R.L. STRATONOVICH

TOPICS IN THE THEORY OF RANDOM NOISE

1

R.L. STRATONOVICH

TOPICS IN THE THEORY OF RANDOM NOISE

1

General Theory of Random Processes Nonlinear Transformations of Signals and Noise

Author's Preface to Volumes I and II

Random processes are of great and ever increasing interest to scientists and engineers working in various branches of radio physics. This is quite understandable, since at the present stage of technological development, random noise constitutes the chief obstacle to further improvement of certain engineering devices. In fact, a current problem in many fields of radio physics is that of evolving techniques for analysis and design of equipment which is immune to noise. Convincing proof of the importance and urgency of this subject is afforded by the large number of books devoted to it, published both in the U.S.S.R. and elsewhere.

The present book, in two volumes, is concerned with a variety of topics which have been left to one side, as it were, but which are nevertheless of great practical importance. In particular, we have in mind problems involving the effects of noise on electronic relays and vacuum-tube oscillators. Recently, there has been increased interest in the operation of switching devices and self-excited systems, not only because of the development of methods of coherent radar detection, but also in connection with such problems as increasing the accuracy of time measurements, designing particle accelerators and computing machines, etc. However, the journal literature in these new fields of statistical radio engineering is hard to penetrate. In the present book, noise in relays and oscillators is studied in an organized fashion, from a rather general point of view. Material that can be found in the literature has been systematized, and the results of some original investigations by the author have been included. The physical treatment of the results obtained may be a bit sketchy, but it seems to us that brevity coupled with wide coverage is justified in this instance, since the rapid growth in the total number of

published papers has increased the need for some kind of survey, even one resembling a handbook.

From a mathematical point of view, the topics studied here involve nonlinear transformations of random functions, mostly transformations with memory. Certain mathematical methods are worked out in a form suitable for application to problems of radio engineering, and are then exploited systematically. As the reader will discover, the following methods turn out to be effective:

1. Reduction of a given process to a process without aftereffect, followed by application of Markov process techniques, in particular, the Fokker-Planck equation;
2. Linearization of the original equations, which allows us to apply correlation theory in the linear approximation;
3. The quasi-static approach, which allows us to reduce a nonlinear transformation with memory to one without memory.

The treatments in the available literature usually deal only with the correlation theory, or only with the theory of Markov processes. However, to solve concrete problems, one must be able to go from the methods of either theory to those of the other, choosing the best methods for solving the problem at hand. For this reason, we pay special attention to the problem of when, and in what sense, the theory of Markov processes can be applied to actual fluctuations, which are originally described in terms of the correlation theory.

Part 1 of Volume I is devoted to a review of the mathematical results used later in the book. In many places, it has seemed appropriate, for the sake of brevity, to depart from the usual presentation of the basic concepts and results of probability theory. The reader who is not familiar with this necessary background material can refer to any one of a number of textbooks. Those already acquainted with elementary probability theory will find that Part 1 contains a different treatment of certain well-known topics, in addition to some new results. In fact, Part 1 ought not to be devoid of interest even to mathematicians, provided, of course, that they make due allowances for the heuristic level of rigor adopted.

In Volume I, Part 2 and in Volume II, we systematically use the methods just enumerated to solve specific problems of statistical radio engineering. At the same time, we try to explain the conditions under which the various methods of analyzing noise phenomena are applicable. Perhaps some of the methods are given in more detail than is necessary for immediate applications, but this is done with a view to further development of the tools needed for future work. It is reasonable to expect that the progressively more complicated problems which come up in engineering practice will require the use of progressively more sophisticated methods of analysis. The author will regard his task as accomplished if this book serves in some measure to broaden the theoretical outlook of radio scientists, by helping them master certain methods which they may find novel at first. Although the material given here is intended primarily for specialists in the field of statistical radio engineering, it should also interest scientists working in other fields where similar statistical methods can be profitably used.

As a rule, no attempt is made to cite the papers where various results appear for the first time. This is because any attempt to make systematic source references would be likely to entail errors, unless one wishes to become involved in special bibliographic research.

Finally, the author would like to take this opportunity to thank V. I. Tikhonov, who suggested that the present book be written and rendered substantial assistance, Y. L. Klimontovich, who in his capacity as editor made many valuable suggestions while the manuscript was being prepared for the printer, and I. G. Akopyan, who constructed the figures appearing in Volume II, Chapter 9. The author would also like to thank (among others) S. P. Strelkov, S. D. Gvozdover, B. R. Levin, S. A. Akhmanov, Y. M. Romanovski and P. S. Lande, for their help with the manuscript. He also expresses his gratitude to Dr. R. A. Silverman for undertaking to translate and edit the English-language edition of this book, thereby greatly enlarging its prospective audience.

R. L. S.

Translator's Preface to Volume I

The present book is the first of two volumes dealing with various topics in the theory of random noise. This material is for the most part not available elsewhere in the English-language literature, and is in fact mainly due to Dr. Stratonovich and his co-workers. Speaking as one who has long been interested in noise theory, I feel confident that a large audience of scientists and engineers will find the contents of this volume interesting, timely, and most important of all, highly original.

I have worked through all the mathematical derivations in the book, and in the process have detected and corrected some typographical errors which had survived in the Russian edition. Various small improvements that occurred to me as I studied the book have been incorporated without special comment. In reply to my queries, Dr. Stratonovich has cleared up a few stubborn points that I was unable to resolve by myself, and the present edition benefits from these clarifications. For example, the supplement to Chapter 4, giving the details of a rather tricky argument, stems from a letter by the author.

Passages appearing between boldface asterisks were printed in brevier originally. They relate to certain side issues which can be omitted without loss of continuity. Finally, bearing in mind the needs of its readers, I have suitably modified the list of books cited in the Bibliography.

R. A. S.

Contents

PART 1 — GENERAL THEORY OF RANDOM PROCESSES

Chapter 1. Random Functions and Their Statistical Characteristics

1. Random Variables 3
2. Correlation Between Random Variables 10
3. Random Functions 15

Chapter 2. Stationary Random Processes and Spectral Densities

1. Basic Concepts 21
2. Calculation of the Spectral Density by Using the Laplace Transform 26
3. The Random Spectrum and its Relation to the Spectral Density 27
4. Time Averages of Stationary Processes. 31
5. Cross-Correlation Functions 35

Chapter 3. Gaussian and Non-Gaussian Random Processes. Quasi-Moment Functions

1. Gaussian Processes and Their Probability Densities 39
2. Conditional Correlation Functions of Gaussian Processes 44
3. Non-Gaussian Processes and Quasi-Moment Functions 47

Chapter 4. Markov Processes and Related Processes

1. Definition of a Markov Process. The Stochastic Equation 55
2. Delta-Correlated Processes 60

3. The Fokker-Planck Equation and the Kolmogorov Equation 62
4. Solutions of the Fokker-Planck Equation 66
5. The Multidimensional Fokker-Planck Equation . . 75
6. First-Passage Time Problems 79
7. Replacement of an Actual Process by a Markov Process. A Special Case 83
8. Replacement of an Actual Process by a Markov Process. The General Case 89
9. Systems of Fluctuation Equations Involving Several Processes 100
10. Gaussian Markov Processes 103
11. Second-Order Fluctuation Equations. Solutions of Special Cases of the Two-Dimensional Fokker-Planck Equation 107
12. Transition from a Markov Process to a Smoothed Process 122
Supplement 126

Chapter 5. Nonstationary Random Processes

1. Processes with Slow Nonstationary Changes . . . 131
2. Switched-On Processes 132
3. Processes with Stationary Increments 133
4. Periodic Nonstationary Processes 139

Chapter 6. Systems of Random Points and Related Random Functions

1. Methods for Describing Systems of Random Points 143
2. Random Functions Constructed from Systems of Random Points 150
3. Poisson Systems of Random Points. Shot Noise . . 153
4. Systems of Random Points Defined by the First Two Moments 159
5. Spectral Densities of Various Pulse Sequences . . 161

Chapter 7. Narrow-Band Random Processes

1. Equivalence of a Narrow-Band Process to a Pair of Slowly Varying Processes 177

2. Narrow-Band Processes Defined by Differential Equations 184
3. The Amplitude and Phase of a Narrow-Band Process. Rayleigh Processes 191
4. Quasi-Rayleigh Processes 198

PART 2 — NONLINEAR TRANSFORMATIONS OF SIGNALS AND NOISE

Chapter 8. Zero-Memory Nonlinear Transformations

1. Transformations of Probability Densities 206
2. Moment Functions for Polynomial Transfer Characteristics 209
3. Moment Functions for Piecewise Linear Transfer Characteristics. The Direct Method 219
4. Moment Functions for Exponential Transfer Characteristics. Rice's Method 233

Chapter 9. Nonlinear Transformations with Memory. Detection of Random Signals

1. Detection of Narrow-Band Processes. The Envelope Method 245
2. The Method of Small Nonlinearity 256
3. The Method of the Fokker-Planck Equation . . . 263
4. Detection of a Sine Wave Plus Noise with a Small Correlation Time 267
5. The Case of Intermediate Correlation Times . . . 271

Bibliography 287

Index 289

PART 1

General Theory of Random Processes

CHAPTER 1

Random Functions and Their Statistical Characteristics

1. Random Variables

Roughly speaking, a *random variable* ξ is characterized by the fact that instead of knowing its precise value, we only know how to obtain various values of ξ under certain experimental conditions which are fixed, stable and physically well-defined. As a result of unknown factors, in a given experiment ξ takes a particular value ξ_1, which we call a *realization* (or *sample value*) of the random variable ξ. For example, consider the number of points which appear when a die is tossed, or the reading of a meter in a certain apparatus at a given instant of time. By repeating the experiment many times (and this possibility is implicit in the definition of a random variable), we can always obtain a large number of realizations. Therefore, a random variable can also be characterized by a certain *statistical ensemble* of realizations. By subjecting the realizations of a given random variable to statistical data processing, we can find certain of its statistical characteristics. Therefore, when we say that a random variable is "given," we mean that we know the statistical data which completely characterize it.

The simplest statistical characteristic of a random variable is its *mean value*: If $\xi_1, \xi_2, \ldots$ are realizations of the random variable ξ, then the mean value of ξ can be defined by

$$\langle \xi \rangle = \lim_{n \to \infty} \frac{\xi_1 + \xi_2 + \cdots + \xi_n}{n}, \tag{1.1}$$

i.e., $\langle \xi \rangle$ is the limit of the arithmetic mean of the sample values as the number of sample values (with respect to which the averaging is performed) is increased without limit. Probability theory can only be used to study experimental data for which such limits exist and do not depend on how the realizations $\xi_1, \xi_2, \ldots, \xi_n$ are chosen from the total statistical ensemble. Naturally, in making this choice, no information involving the specific values of the realizations should be used.[1]

Other statistical characteristics of a random variable ξ can be defined by first going over to a new random variable $\eta = f(\xi)$, and then using formula (1.1) to calculate $\langle \eta \rangle$; in so doing, we explicitly use the form of the function $f(\xi)$. Thus, setting $f(\xi) = \xi^2$, we define the *mean square* of ξ by the formula

$$\langle \xi^2 \rangle = \lim_{n \to \infty} \frac{\xi_1^2 + \xi_2^2 + \ldots + \xi_n^2}{n}. \tag{1.2}$$

As another example, consider the probability $\mathbf{P}\{\xi < x\}$ that the inequality $\xi < x$ holds, where x is a fixed number. To calculate $\mathbf{P}\{\xi < x\}$, we introduce the random variable

$$\epsilon = \vartheta(x - \xi), \quad \text{where} \quad \vartheta(z) = \begin{cases} 1 & \text{for} \quad z > 0, \\ 0 & \text{for} \quad z \leqslant 0. \end{cases} \tag{1.3}$$

Then we have

$$\mathbf{P}\{\xi < x\} = \langle \vartheta(x - \xi) \rangle,$$

[1] We address the following remark to the reader who is already acquainted with the subject: The approach based on the limit (1.1) is simple, intuitive and quite adequate for use in the applications; it has certain defects from a formal mathematical point of view, but these are unimportant when dealing with applications of probability theory. The point is that as far as the applications are concerned, the assertion that a given event has probability equal to unity (or arbitrarily close to unity) is equivalent to the assertion that the event occurs. Therefore, instead of the familiar requirement that the convergence expressed by (1.1) occurs with probability arbitrarily close to unity, we choose as our starting point the requirement that (1.1) actually converges. The reader will find other ways of dealing with this point in other treatments of the subject.

i.e.,

$$\mathbf{P}\{\xi < x\} = \lim_{n\to\infty} \frac{\epsilon_1 + \epsilon_2 + \ldots + \epsilon_n}{n}. \tag{1.4}$$

In fact, bearing in mind the definition of the function $\vartheta(z)$, we see that the average of the arithmetic mean $(\epsilon_1 + \epsilon_2 + \ldots + \epsilon_n)/n$ is just the ratio m/n, where m is the number of inequalities $\xi_i < x$ $(1 \leqslant i \leqslant n)$ which are valid, and n is the total number of trials. Thus, (1.4) can be written in the familiar form

$$\mathbf{P}\{\xi < x\} = \lim_{n\to\infty} \frac{m}{n}. \tag{1.5}$$

The mean value $\langle \vartheta(x - \xi)\rangle$ regarded as a function of x, i.e.,

$$F_\xi(x) = \langle \vartheta(x - \xi)\rangle, \tag{1.6}$$

is called the *distribution function* of ξ, and its derivative

$$w_\xi(x) = \frac{d}{dx} F_\xi(x) \qquad \left(F_\xi(x) = \int_{-\infty}^{x} w_\xi(x')\, dx'\right) \tag{1.7}$$

is called the *probability density* of ξ. Substituting (1.6) into (1.7), and formally differentiating under the averaging sign, we obtain

$$w_\xi(x) = \langle \delta(\xi - x)\rangle, \tag{1.8}$$

where

$$\delta(z) = \frac{d}{dz}\vartheta(z)$$

is the *Dirac delta function*, which has the property that

$$\int_{-\infty}^{\infty} f(z)\,\delta(z - z_0)\, dz = f(z_0) \quad \text{for any function } f(z).$$

In what follows, we shall often be concerned with probability densities. For brevity, it is sometimes convenient to denote the value of x by the same letter as the random variable ξ itself, so that $w_x(\xi)$ becomes $w_\xi(\xi)$. We then drop the redundant subscript ξ and write simply $w(\xi)$ for the probability density of the random variable ξ.

The probability density $w(\xi)$ satisfies the *normalization condition*[2]

$$\int w(\xi)\, d\xi = 1\,, \tag{1.9}$$

which, according to (1.8), is equivalent to the trivial identity

$$\langle 1 \rangle = 1\,.$$

A random variable ξ is completely characterized by its probability density, and knowing this probability density, we can calculate the mean value $\langle f(\xi) \rangle$ of any function $f(\xi)$. In fact, using the basic property of the delta function, we can first write $f(\xi)$ as

$$f(\xi) = \int f(x)\, \delta(\xi - x)\, dx$$

and then calculate $\langle f(\xi) \rangle$. Taking the average under the integration sign, we see at once that

$$\langle f(\xi) \rangle = \int f(x)\, \langle \delta(\xi - x) \rangle\, dx\,,$$

or

$$\langle f(\xi) \rangle = \int f(x)\, w_\xi(x)\, dx\,, \tag{1.10}$$

because of (1.8).

Next, we show how the probability density behaves under transformations of the original random variable ξ. Let the new random variable η be defined by

$$\eta = g(\xi)\,, \tag{1.11}$$

where $g(x)$ is a known function. Applying (1.8), we write the required probability density of the new random variable η in the form

$$w_\eta(y) = \langle \delta(\eta - y) \rangle = \langle \delta[g(\xi) - y] \rangle\,.$$

---

[2] When limits of integration are omitted, they will always be understood to be from $-\infty$ to ∞.

It then follows from (1.10) that

$$w_\eta(y) = \int \delta[g(x) - y]\, w_\xi(x)\, dx\,, \tag{1.12}$$

or more concisely,

$$w(\eta) = \int \delta[g(\xi) - \eta]\, w(\xi)\, d\xi\,. \tag{1.13}$$

This last integral can be evaluated by using properties of the delta function. If the function (1.11) is monotone, it has a unique inverse function

$$\xi = h(\eta) \qquad (h = g^{-1})\,. \tag{1.14}$$

In this case, it follows from (1.12) that

$$w(\eta) = \frac{w(\xi)}{\left|\dfrac{dg(\xi)}{d\xi}\right|_{\xi=h(\eta)}} = w[h(\eta)] \left|\frac{d\xi}{d\eta}\right|_{\xi=h(\eta)}. \tag{1.15}$$

In the general case where the function $g(\xi)$ is not monotone, and instead has "several inverse functions"

$$\xi = h_l(\eta) \qquad (l = 1, 2, \ldots, r)\,, \tag{1.16}$$

formula (1.12) gives

$$w(\eta) = \sum_{l=1}^{r} w[h_l(\eta)] \left|\frac{d\xi}{d\eta}\right|_{\xi=h_l(\eta)}. \tag{1.17}$$

Besides the probability density $w(\xi)$, the *characteristic function*

$$\Theta_\xi(u) = \langle e^{iu\xi} \rangle = \int e^{iu\xi} w(\xi)\, d\xi \tag{1.18}$$

completely characterizes the random variable ξ. In fact, $w(\xi)$ is just the Fourier transform of $\Theta_\xi(u)$:

$$w(\xi) = \frac{1}{2\pi} \int e^{-iu\xi}\, \Theta_\xi(u)\, du\,. \tag{1.19}$$

The *moments*

$$m_n = \langle \xi^n \rangle \qquad (n = 1, 2, \ldots) \tag{1.20}$$

of the random variable ξ can be obtained from the characteristic function by differentiation

$$m_n = \frac{1}{i^n} \frac{d^n \Theta_\xi(u)}{du^n}\bigg|_{u=0}. \tag{1.21}$$

From a knowledge of the moments m_n, i.e., the values of the derivatives of the characteristic function $\Theta_\xi(u)$ at the origin, we can write $\Theta_\xi(u)$ as a Maclaurin series:

$$\Theta_\xi(u) = 1 + \sum_{n=1}^{\infty} \frac{(iu)^n}{n!} m_n . \tag{1.22}$$

Of course, we can only write (1.22) when the moments m_n are finite and do not grow too rapidly as the index n is increased.

To completely describe a random variable, one has to know all its moments m_1, m_2, ..., but most of the physically important information is contained in the first few moments. The higher moments describe rather unimportant details in the behavior of the probability density, which pertain mostly to extreme values of the random variable. For a variety of reasons, it is more convenient to characterize a random variable not by its moments, but by its *cumulants* (or *semi-invariants*) k_n, defined by the relation

$$\Theta_\xi(u) = \exp\left\{\sum_{n=1}^{\infty} \frac{(iu)^n}{n!} k_n\right\}. \tag{1.23}$$

Comparing (1.22) and (1.23), we can obtain the following formulas relating the moments and the cumulants:

$$\begin{aligned} k_1 &= m_1 , \\ k_2 &= m_2 - m_1^2 , \\ k_3 &= m_3 - 3m_1 m_2 + 2m_1^3 , \\ k_4 &= m_4 - 3m_2^2 - 4m_1 m_3 + 12 m_1^2 m_2 - 6m_1^4 , \\ &\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned} \tag{1.24}$$

From the structure of these formulas, it is clear that if we know the first n moments, we can calculate the first n cumulants, and conversely, i.e., the first n cumulants contain just as much information as the first n moments. If we confine ourselves to a definite number of first cumulants, it is convenient to set the higher cumulants equal to zero. To deal with moments in the same way, i.e., to set all moments higher than a certain order equal to zero, is not desirable, since then the characteristic function would behave at infinity in a way which makes it difficult to calculate the inversion integral (1.19). In fact, $w(\xi)$ would then become a sum of derivatives of delta functions, with no simple intuitive meaning.

The first cumulant is the same as the first moment, i.e., $m_1 = k_1 = \langle\xi\rangle$, and is just the mean value, the most important characteristic of a random variable. The next two cumulants are the central moments of orders two and three:

$$k_2 = \langle(\xi - \langle\xi\rangle)^2\rangle\,, \qquad k_3 = \langle(\xi - \langle\xi\rangle)^3\rangle\,. \tag{1.25}$$

The quantity $k_2 = \langle\xi^2\rangle - \langle\xi\rangle^2$ is called the *variance* (or *dispersion*) of ξ, and is denoted by

$$\mathbf{D}\xi \equiv \langle\xi^2\rangle - \langle\xi\rangle^2\,. \tag{1.26}$$

The variance is also a very important characteristic of a random variable, since the smaller the variance, the more exact our knowledge of the random variable. The probability that the absolute value of the deviation of ξ from its mean value $\langle\xi\rangle$ exceeds a fixed $\epsilon > 0$ approaches zero as $\mathbf{D}\xi$ approaches zero. Therefore, if $\mathbf{D}\xi$ is small enough, we can assume that ξ differs by very little from $\langle\xi\rangle$.

Instead of the variance, one often deals with the *standard deviation*

$$\sigma(\xi) \equiv [\mathbf{D}\xi]^{1/2} = [\langle\xi^2\rangle - \langle\xi\rangle^2]^{1/2}\,. \tag{1.27}$$

While retaining the important features of the variance, the standard deviation has the advantage that it has the same physical dimension as the random variable itself. The standard deviation indicates the

order of magnitude of random deviations from the mean value, i.e.,

$$\xi - \langle \xi \rangle \sim \sigma(\xi) . \tag{1.28}$$

For the time being, we shall not discuss the significance of higher cumulants. However, we note that the ratios

$$\gamma_a = \frac{k_3}{k_2^{3/2}}, \qquad \gamma_e = \frac{k_4}{k_2^2} \tag{1.29}$$

are called the *coefficient of asymmetry* and the *coefficient of excess*, respectively. The physical meaning of the cumulant k_n decreases as n increases.

2. Correlation Between Random Variables

Suppose we have r random variables $\xi_1, \ldots, \xi_r$. Loosely speaking, this means that as a result of each experiment (trial) we simultaneously obtain a realization of each of the r random variables. The random variables $\xi_1, \ldots, \xi_r$ are completely characterized by the (*joint*) *r-dimensional probability density*

$$w_{\xi_1 \ldots \xi_r}(x_1, \ldots, x_r) = \langle \delta(\xi_1 - x_1) \ldots \delta(\xi_r - x_r) \rangle , \tag{1.30}$$

or simply $w(\xi_1, \ldots, \xi_r)$, which is a function of r variables. It follows by definition that the mean value of any function $f(\xi_1, \ldots, \xi_r)$ can be calculated by integrating with weight $w(\xi_1, \ldots, \xi_r)$:

$$\langle f(\xi_1, \ldots, \xi_r) \rangle = \int \ldots \int f(\xi_1, \ldots, \xi_r)\, w(\xi_1, \ldots, \xi_r)\, d\xi_1 \ldots d\xi_r . \tag{1.31}$$

If we set $f \equiv 1$, we obtain the normalization condition

$$\int \ldots \int w(\xi_1, \ldots, \xi_r)\, d\xi_1 \ldots d\xi_r = 1 . \tag{1.32}$$

Next, we divide the r random variables $\xi_1, \ldots, \xi_r$ into two groups $\xi_1, \ldots, \xi_k$ and $\xi_{k+1}, \ldots, \xi_r$, and we ask whether knowing the actual realizations of the random variables $\xi_{k+1}, \ldots, \xi_r$ gives us any

information about $\xi_1, \ldots, \xi_k$. For example, suppose we have two random variables ξ_1 and ξ_2, described by a probability density $w(\xi_1, \xi_2)$. In the absence of knowledge of the actual realization (i.e., observed value) of ξ_2, the random variable ξ_1 has the probability density

$$w(\xi_1) = \int w(\xi_1, \xi_2)\, d\xi_2 . \tag{1.33}$$

Knowing the observed value of the random variable ξ_2 (which gives us complete information about ξ_2), we can also obtain some information about the other random variable ξ_1. Of course, in general our knowledge of ξ_1 is still imperfect, even when we know the observed value of ξ_2, and hence ξ_1 remains undetermined, i.e., is still a random variable. However, ξ_1 is now described by another probability density

$$w(\xi_1 \mid \xi_2) = \frac{w(\xi_1, \xi_2)}{\int w(\xi_1, \xi_2)\, d\xi_1} = \frac{w(\xi_1, \xi_2)}{w(\xi_2)}, \tag{1.34}$$

which contains more (or at least not less) information about ξ_1 than the distribution (1.33).[3] The expression (1.34) is called the *conditional probability density* of ξ_1, given the value of ξ_2. The integral appearing in the denominator of the middle term in (1.34) guarantees that the normalization condition

$$\int w(\xi_1 \mid \xi_2)\, d\xi_1 = 1 \tag{1.35}$$

is satisfied.

Just how much information is gained about ξ_1 by knowing the value of ξ_2 depends on the particular case. In some cases, no information about ξ_1 is gained, regardless of the observed value of ξ_2, which means that

$$w(\xi_1 \mid \xi_2) = w(\xi_1) . \tag{1.36}$$

[3] As is customary, probability densities will often be referred to (somewhat imprecisely) as "distributions."

We see from (1.34) that in this case $w(\xi_1, \xi_2)$ is just the product of the two one-dimensional probability densities $w(\xi_1)$ and $w(\xi_2)$, i.e.,

$$w(\xi_1, \xi_2) = w(\xi_1)\, w(\xi_2)\,, \tag{1.37}$$

and the random variables ξ_1 and ξ_2 are said to be (*statistically*) *independent*. The opposite extreme is the case where knowledge of ξ_2 completely determines the other random variable ξ_1, which is a function of ξ_2 [$\xi_1 = f(\xi_2)$]. Then, ξ_1 and ξ_2 are said to be *completely correlated*, and their joint distribution contains a delta function:

$$w(\xi_1, \xi_2) = \delta[\xi_1 - f(\xi_2)]\, w(\xi_2)\,. \tag{1.38}$$

Naturally, there is a whole range of intermediate cases between these two extreme cases.

All that has been said above about two random variables can immediately be generalized to the case of several random variables. Given r random variables $\xi_1, \ldots, \xi_r$, knowledge of several of them, e.g., $\xi_{k+1}, \ldots, \xi_r$ increases (or at least does not decrease) our information about the other random variables $\xi_1, \ldots, \xi_k$. We can express the conditional probability density of $\xi_1, \ldots, \xi_k$ in terms of the original probability density $w(\xi_1, \ldots, \xi_r)$ as follows:

$$w(\xi_1, \ldots, \xi_k \mid \xi_{k+1}, \ldots, \xi_r) = \frac{w(\xi_1, \ldots, \xi_r)}{w(\xi_{k+1}, \ldots, \xi_r)}\,. \tag{1.39}$$

If knowledge of $\xi_{k+1}, \ldots, \xi_r$ does not increase our information about $\xi_1, \ldots, \xi_k$, then the two groups of random variables $\xi_1, \ldots, \xi_k$ and $\xi_{k+1}, \ldots, \xi_r$ are said to be (*statistically*) *independent*. However, if the conditional distribution (1.39) does not coincide with the "unconditional" distribution

$$w(\xi_1, \ldots, \xi_k) = \int \cdots \int w(\xi_1, \ldots, \xi_r)\, d\xi_{k+1} \cdots d\xi_r\,, \tag{1.40}$$

then there is statistical dependence between the two groups of random variables.

It is now convenient to introduce numerical characteristics which describe the degree of statistical dependence between random variables, and which vanish when there is no statistical dependence

between them. Such characteristics are provided by *multiple correlations* of various orders. The *covariance* (synonymously, *cross correlation* or *double correlation*) of two random variables ξ_1 and ξ_2, denoted by $\mathbf{K}[\xi_1, \xi_2]$, is defined by the relation

$$\mathbf{K}[\xi_1, \xi_2] = \langle \xi_1 \xi_2 \rangle - \langle \xi_1 \rangle \langle \xi_2 \rangle . \tag{1.41}$$

If ξ_1 and ξ_2 are independent, we have

$$\langle \xi_1 \xi_2 \rangle = \langle \xi_1 \rangle \langle \xi_2 \rangle ,$$

according to (1.37). Hence, in this case, the covariance vanishes. By definition, the covariance of a random variable with itself is just its variance

$$\mathbf{K}[\xi, \xi] = \mathbf{D}\xi . \tag{1.42}$$

Sometimes, instead of $\mathbf{K}[\xi_1, \xi_2]$, it is convenient to consider the dimensionless quantity

$$R_2 \equiv R = \frac{\mathbf{K}[\xi_1, \xi_2]}{\sigma(\xi_1)\, \sigma(\xi_2)} , \tag{1.43}$$

which is proportional to $\mathbf{K}[\xi_1, \xi_2]$. The quantity (1.43) reduces to unity when the random variables coincide, and is called the *correlation coefficient* of ξ_1 and ξ_2.

Besides double correlations, there are multiple correlations of higher orders. Thus, given three random variables ξ_1, ξ_2, ξ_3, we can construct a *triple correlation* $\mathbf{K}[\xi_1, \xi_2, \xi_3]$, which describes statistical dependence involving all three random variables ξ_1, ξ_2, ξ_3, and vanishes when at least one of the random variables is independent of the others. We first form the quantity

$$\langle \xi_1 \xi_2 \xi_3 \rangle - \langle \xi_1 \rangle \langle \xi_2 \rangle \langle \xi_3 \rangle , \tag{1.44}$$

by subtracting the product of the mean values of the random variables from their joint third moment. The quantity (1.44) has the property that it vanishes when all three random variables are independent of each other; however, (1.44) can be different from zero not only because of correlations involving all three random variables, but also because of correlations between pairs of random

variables. For example, suppose that one of the random variables, ξ_3 say, is independent of the other two. Then, there is no triple correlation, but the quantity (1.44) can be different from zero because of the presence of double correlation between ξ_1 and ξ_2. In fact, in this case, we have $\langle \xi_1 \xi_2 \xi_3 \rangle = \langle \xi_1 \xi_2 \rangle \langle \xi_3 \rangle$, and (1.44) equals

$$\langle \xi_3 \rangle \mathbf{K}[\xi_1, \xi_2] . \tag{1.45}$$

Thus, it is clear that to properly define the triple correlation $\mathbf{K}[\xi_1, \xi_2, \xi_3]$, we have to subtract from (1.44) all terms of the form (1.45) due to the presence of double correlations. There are three distinct pairs which can be selected from ξ_1, ξ_2, ξ_3, i.e.,

$$\xi_1, \xi_2 ; \qquad \xi_2, \xi_3 ; \qquad \xi_1, \xi_3 .$$

Therefore, subtracting all terms involving double correlations from (1.44), we define the *triple correlation* $\mathbf{K}[\xi_1, \xi_2, \xi_3]$ by the formula

$$\begin{aligned} \mathbf{K}[\xi_1, \xi_2, \xi_3] = \langle \xi_1 \xi_2 \xi_3 \rangle &- \langle \xi_1 \rangle \mathbf{K}[\xi_2, \xi_3] - \langle \xi_2 \rangle \mathbf{K}[\xi_1, \xi_3] \\ &- \langle \xi_3 \rangle \mathbf{K}[\xi_1, \xi_2] - \langle \xi_1 \rangle \langle \xi_2 \rangle \langle \xi_3 \rangle . \end{aligned} \tag{1.46}$$

Using (1.41), we can write (1.46) in the form

$$\begin{aligned} \mathbf{K}[\xi_1, \xi_2, \xi_3] = \langle \xi_1 \xi_2 \xi_3 \rangle &- \langle \xi_1 \rangle \langle \xi_2 \xi_3 \rangle - \langle \xi_2 \rangle \langle \xi_1 \xi_3 \rangle \\ &- \langle \xi_3 \rangle \langle \xi_1 \xi_2 \rangle + 2 \langle \xi_1 \rangle \langle \xi_2 \rangle \langle \xi_3 \rangle . \end{aligned} \tag{1.47}$$

In the case where all three random variables coincide, the triple correlations (1.47) reduces to the third cumulant

$$\mathbf{K}[\xi, \xi, \xi] = k_3 . \tag{1.48}$$

Moreover, by analogy with (1.43), we can introduce the *triple correlation coefficient*

$$R_3 = \mathbf{K}[\xi_1, \xi_2, \xi_3] \{ \mathbf{K}[\xi_1, \xi_1, \xi_1] \mathbf{K}[\xi_2, \xi_2, \xi_2] \mathbf{K}[\xi_3, \xi_3, \xi_3] \}^{-1/3} . \tag{1.49}$$

To define the *fourth-order correlation* for four random variables, we have to subtract all possible double and triple correlations from the difference

$$\langle \xi_1 \xi_2 \xi_3 \xi_4 \rangle - \langle \xi_1 \rangle \langle \xi_2 \rangle \langle \xi_3 \rangle \langle \xi_4 \rangle . \tag{1.50}$$

The number of terms to be subtracted is determined by the number of possible combinations of the elements $\xi_1, \xi_2, \xi_3, \xi_4$ arranged in groups of one, two or three elements. The formulas for higher-order correlations become more complicated as the order is increased.

The most concise expression for mean values of products of several random variables can be written in terms of the characteristic function

$$\Theta(u_1, \ldots, u_r) = \langle \exp i(u_1\xi_1 + \ldots + u_r\xi_r) \rangle . \tag{1.51}$$

In fact, we have

$$\langle \xi_1 \ldots \xi_r \rangle = \frac{1}{i^r} \frac{\partial^r \Theta(u_1, \ldots, u_r)}{\partial u_1 \ldots \partial u_r} \Big|_{u_1=\ldots=u_r=0} . \tag{1.52}$$

To obtain the corresponding formula for the multiple correlation of order r, we need only replace the characteristic function by its logarithm in (1.52):

$$\mathbf{K}[\xi_1, \ldots, \xi_r] = \frac{1}{i^r} \frac{\partial^r \ln \Theta(u_1, \ldots, u_r)}{\partial u_1 \ldots \partial u_r} \Big|_{u_1=\ldots=u_r=0} . \tag{1.53}$$

3. Random Functions

Next, we consider a *random function* $\xi(t)$ of a single real argument t (the time) which varies over the interval $[0, T]$, say. The random function $\xi(t)$ will be regarded as given if we know stable experimental conditions which allow us to obtain realizations of $\xi(t)$ in the interval $[0, T]$. For example, such a realization might be a photograph (or oscillogram) of the random function, whose form depends on time.

If we take r fixed values $t_1, \ldots t_r$ from the interval $[0, T]$, where r, $t_1, \ldots, t_r$ are arbitrary, then the values $\xi(t_1), \ldots, \xi(t_r)$ constitute a family of random variables. If the random function is completely specified, these random variables are also completely specified, for arbitrary $r, t_1, \ldots, t_r$. Therefore, we can regard the corresponding

probability density $w[\xi(t_1), ..., \xi(t_r)]$ as specified. This probability density is a function of r, $t_1, ..., t_r$, and can be written as

$$\langle \delta[x_1 - \xi(t_1)] \dots \delta[x_r - \xi(t_r)] \rangle = w_r(x_1, ..., x_r; t_1, ..., t_r) . \tag{1.54}$$

Each probability density (1.54), regarded as a function of $x_1, ..., x_r$, satisfies the usual requirements imposed on probability densities, i.e., nonnegativity and the normalization condition. At the same time, as functions of $t_1, ..., t_r$, the probability densities (1.54) satisfy certain additional conditions, which are a consequence of its definition. These are the *symmetry condition*

$$w_r(..., x_k, ..., x_l, ...; ..., t_k, ..., t_l, ...) = w_r(..., x_l, ..., x_k, ...; ..., t_l, ..., t_k, ...), \tag{1.55}$$

and the *compatibility condition*

$$w_r(x_1, ..., x_r; t_1, ..., t_r) = \int w_{r+k}(x_1, ..., x_{r+k}; t_1, ..., t_{r+k})\, dx_{r+1} \dots dx_{r+k} . \tag{1.56}$$

We can consider a random function $\xi(t)$ to be specified if the probability densities (1.54) are specified for all r, no matter how large. By choosing the points of division $t_1, t_2, ..., t_r$ sufficiently close together, we can replace $\xi(t)$ by a sequence of random variables $\xi(t_1), \xi(t_2), ..., \xi(t_r)$, with an accuracy which is sufficient for all practical purposes. Thus, a random function $\xi(t)$ can be characterized by an infinite sequence of probability densities

$$w_1(x_1; t) , \qquad w_2(x_1, x_2; t_1, t_2) , \quad ... , \tag{1.57}$$

which satisfy all the necessary requirements.

If we know the first r probability densities, we can discard the probability densities $w_1, ..., w_{r-1}$, since they contain no extra information not already contained in the probability density $w_r(x_1, ..., x_r; t_1, ..., t_r)$. In this regard, it should be noted that the ideal state of affairs would be to describe a random process by a single probability density of the largest possible order, if such existed, but since the argument t is continuous, there is no probability density of maximum finite order. However, it is

sometimes possible to introduce a *probability functional* $W[\xi(t)]$, which defines a "continuous probability density," i.e., defines a probability density in an appropriate function space.

Instead of the probability densities (1.57), we can give the following sequence of characteristic functions:

$$\begin{aligned} \Theta_1(u_1; t_1) &= \langle e^{iu_1\xi(t_1)} \rangle , \\ \Theta_2(u_1, u_2; t_1, t_2) &= \langle e^{iu_1\xi(t_1)+iu_2\xi(t_2)} \rangle , \\ &\cdots \quad . \quad \cdots \end{aligned} \tag{1.58}$$

Specifying (1.58) is equivalent to specifying the sequence of distributions (1.57). Here, as before, we need only retain the characteristic function of highest order (if there is one), since all characteristic functions of lower orders can be obtained from the one of highest order by setting certain arguments equal to zero:

$$\Theta_r(u_1, \ldots, u_r; t_1, \ldots, t_r) = \Theta_{r+k}(u_1, \ldots, u_r, 0, \ldots, 0; t_1, \ldots, t_{r+k}). \tag{1.59}$$

By increasing the order r, thereby approaching the "most complete characteristic function," containing as much information as possible about $\xi(t)$, we obtain in the limit the *characteristic functional*

$$\Theta[u(t)] = \left\langle \exp\left[i \int u(t)\, \xi(t)\, dt\right] \right\rangle \tag{1.60}$$

which completely describes the random process $\xi(t)$. In this formula, the integral in the exponential extends over the whole domain of definition of the random function. The characteristic functions (1.58) can be obtained from (1.60) by choosing argument functions of the form

$$u(t) = \sum u_\alpha \delta(t - t_\alpha) . \tag{1.61}$$

In its general form, a random process[4] cannot be described by a finite number of functions of a finite number of variables, and hence a random process is either characterized by an infinite

[4] The term *random process* (or simply *process*) is used as a synonym for a random function of time.

sequence of functions or by a functional. Instead of the sequences (1.57) and (1.58), it is preferable to consider a sequence of functions such that the functions of higher order do not repeat the information contained in the preceding functions, but instead carry only new information. Therefore, we now introduce another method of describing random processes, which has this and other advantages.

The mean values

$$\langle \xi(t_1) \dots \xi(t_r) \rangle \equiv m_r(t_1, \dots, t_r), \tag{1.62}$$

regarded as functions of $t_1, \dots, t_r$, are called *moment functions*, and the infinite sequence of moment functions

$$m_1(t_1), \quad m_2(t_1, t_2), \quad m_3(t_1, t_2, t_3), \quad \dots \tag{1.63}$$

gives an exhaustive description of the random process $\xi(t)$. It is even better to characterize the random process $\xi(t)$ by the sequence of *correlation functions*

$$k_1(t_1), \quad k_2(t_1, t_2), \quad k_3(t_1, t_2, t_3), \quad \dots, \tag{1.64}$$

which are defined as the multiple correlations

$$k_r(t_1, \dots, t_r) \equiv \mathbf{K}[\xi(t_1), \dots, \xi(t_r)], \tag{1.65}$$

regarded as functions of the times $t_1, \dots, t_r$. The correlation functions, just like the moment functions, are symmetric functions of their arguments. As the order increases, the physical significance of multiple correlations decreases, and hence the first few functions in the sequence (1.64) are the most important. This is the advantage of describing a random process by its correlation functions. The vast majority of random processes encountered in radio physics have the property that when the times $t_1, \dots, t_r$ are moved apart, the correlation between the corresponding values of the process falls off, i.e., the correlation functions (1.65) go to zero (for $r > 1$), which is very convenient.

If we know the moment functions or the correlation functions, we can find the other characteristics of a random process, in particular, its characteristic functions or probability densities. In

fact, using (1.52) we can write the characteristic function as a multidimensional Taylor's series:

$$\Theta_r(u_1, \ldots, u_r; t_1, \ldots, t_r) = 1 + \sum_{s=1}^{\infty} \frac{i^s}{s!} \sum_{\alpha, \ldots, \omega=1}^{r} m_s(t_\alpha, \ldots, t_\omega)\, u_\alpha \ldots u_\omega . \quad (1.66)$$

In just the same way, using (1.53), we can express the characteristic function in terms of the correlation functions:

$$\Theta_r(u_1, \ldots, u_r; t_1, \ldots, t_r) = \exp\left\{\sum_{s=1}^{\infty} \frac{i^s}{s!} \sum_{\alpha, \ldots, \omega=1}^{r} k_s(t_\alpha, \ldots, t_\omega)\, u_\alpha \ldots u_\omega\right\} . \quad (1.67)$$

It is important to note that (1.67) does not change its form if we choose different times $t_1, \ldots, t_r$ or change the number r. Using this fact, we can write the characteristic functional (1.60) as

$$\Theta[u(t)] = \exp\left\{\sum_{s=1}^{\infty} \frac{i^s}{s!} \int \ldots \int k_s(t_1, \ldots, t_s)\, u(t_1) \ldots u(t_s)\, dt_1 \ldots dt_s\right\}. \quad (1.68)$$

In Chap. 3, we shall show how to go from characteristic functions to probability densities.

Because of the special role played by moment functions and correlation functions, we give the following formulas relating them:

$$\begin{aligned}
m_1(t_1) &= k_1(t_1) , \\
m_2(t_1, t_2) &= k_2(t_1, t_2) + k_1(t_1)\, k_1(t_2) , \\
m_3(t_1, t_2, t_3) &= k_3(t_1, t_2, t_3) + 3\{k_1(t_1)\, k_2(t_2, t_3)\}_s + k_1(t_1)\, k_1(t_2)\, k_1(t_3) , \\
m_4(t_1, t_2, t_3, t_4) &= k_4(t_1, t_2, t_3, t_4) + 3\{k_2(t_1, t_2)\, k_2(t_3, t_4)\}_s \\
&\quad + 4\{k_1(t_1)\, k_3(t_2, t_3, t_4)\}_s + 6\{k_1(t_1)\, k_1(t_2)\, k_2(t_3, t_4)\}_s \\
&\quad + k_1(t_1)\, k_1(t_2)\, k_1(t_3)\, k_1(t_4) , \\
&\ldots \quad . \quad \ldots\ldots\ldots
\end{aligned} \quad (1.69)$$

Here, the symbol $\{\ldots\}_s$ denotes the operation of symmetrizing the expression in brackets with respect to all its arguments. If the function to be symmetrized depends on n arguments and has no symmetry at all, this means that we have to carry out all $n!$ possible

permutations of the arguments and then take the arithmetic mean of the resulting $n!$ terms. However, if the function has symmetry with respect to some of its arguments, then among these $n!$ terms, there will be some that are identical. Therefore, we need not calculate the arithmetic mean of all the permuted terms, but only of those which are different. It is interesting to note that the coefficients appearing before each symmetrized expression in the formulas (1.69) are equal to the total number of permuted terms which are not identically equal. To form the arithmetic mean of the terms which are not identically equal, we have to divide the sum of these terms by a number which just equals the factor standing in front of the sum. Thus, all that remains in (1.69) are sums over those permutations of the arguments which lead to terms that do not coincide.

The system (1.69) can be solved step by step for the correlation functions in terms of the moment functions. It should be noted that the correlation functions are *not* the same as the central moments. The difference begins with k_4. Thus, if we set $k_1 = 0$ in (1.69), we find that m_4, the central moment of order four, differs from k_4 by the term

$$m_4(t_1, t_2, t_3, t_4) - k_4(t_1, t_2, t_3, t_4) = k_2(t_1, t_2)\, k_2(t_3, t_4) + k_2(t_1, t_3)\, k_2(t_2, t_4) + k_2(t_1, t_4)\, k_2(t_2, t_3)\,. \tag{1.70}$$

CHAPTER 2

Stationary Random Processes and Spectral Densities

1. Basic Concepts

A random process is said to be *stationary* (in the strict sense) if its statistical characteristics are invariant under time shifts, i.e., if they remain the same when t is replaced by $t + a$, where a is arbitrary. Then, the probability densities $w_n(\xi_1, ..., \xi_n; t_1, ..., t_n)$, together with the moment and correlation functions $m_n(t_1, ..., t_n)$ and $k_n(t_1, ..., t_n)$ do not depend on the absolute position of the points $t_1, ..., t_n$ on the time axis, but only on their relative configuration, i.e., only on the time differences $t_2 - t_1, ..., t_n - t_1$. For example, consider the correlation function $k_n(t_1, ..., t_n)$. By the definition of stationarity, we have

$$k_n(t_1, ..., t_n) = k_n(t_1 + a, ..., t_n + a) . \tag{2.1}$$

Choosing $a = -t_1$ and writing

$$k_n(0, t_2 - t_1, ..., t_n - t_1) = k'_n(t_2 - t_1, ..., t_n - t_1) ,$$

we find that

$$k_n(t_1, ..., t_n) = k'_n(t_2 - t_1, ..., t_n - t_1) . \tag{2.2}$$

Thus, the correlation functions of a stationary process depend only on the time differences, as asserted. In particular, it follows that the mean value $k_1(t_1)$ of a stationary process [equal to $k_1(0)$] must be a constant, which we denote by m:

$$k_1(t) = m_1(t) = m . \tag{2.3}$$

Similarly, the correlation function $k_2(t_1, t_2)$ is a function of $t_2 - t_1$, for which we introduce the special notation

$$k_2(t_1, t_2) = k(t_2 - t_1), \tag{2.4}$$

where $k(\tau) = k(-\tau)$. We can also write $k(\tau)$ in the form

$$k(\tau) = \sigma^2 R(\tau), \tag{2.5}$$

where $\sigma = \sigma[\xi(t)] = \sqrt{k(0)}$ is the standard deviation of $\xi(t)$, and $R(\tau) = k(\tau)/\sigma^2$ is the dimensionless, normalized correlation coefficient (1.43), with the property that $R(0) = 1$. We shall not introduce any special notation for the higher-order correlation functions, since they are not as important. It should be noted that when all the arguments of the third-order and fourth-order correlation functions coincide, we obtain the cumulants k_3 and k_4, appearing in the definition (1.29) of the coefficients of asymmetry and excess:

$$\gamma_a = \sigma^{-3} k_3'(0, 0), \qquad \gamma_e = \sigma^{-4} k_4'(0, 0, 0). \tag{2.6}$$

In what follows, we shall often use the concept of the *correlation time*, defined by the relation

$$\tau_{cor} = \int_0^\infty | R(\tau) | \, d\tau = \sigma^{-2} \int_0^\infty | k(\tau) | \, d\tau. \tag{2.7}$$

The correlation time τ_{cor} gives some idea of the size of the time interval over which correlation extends between values of the process $\xi(t)$. In fact, we can neglect correlation between values of the process whose time separation is considerably greater than τ_{cor}. We also define the quantity

$$K = \int_{-\infty}^\infty k(\tau) \, d\tau, \tag{2.8}$$

which we call the *intensity coefficient* of the random process $\xi(t)$. This quantity plays an important role in the theory of Markov processes.

Next, we introduce the important concept of the (*power*) *spectral density* $S[\xi; \omega]$ of a stationary random process $\xi(t)$:

$$S[\xi; \omega] = 2\int_{-\infty}^{\infty} e^{i\omega\tau}\langle\xi\xi_\tau\rangle\, d\tau = 4\int_0^{\infty} \cos\omega\tau\langle\xi\xi_\tau\rangle\, d\tau\,. \tag{2.9}$$

Here, and subsequently, the subscript τ on a function like ξ_τ, in the presence of a similar function like ξ without a subscript, denotes a shift of the argument by an amount τ, i.e., $\xi_\tau = \xi(t + \tau)$, while $\xi = \xi(t)$. For a process with zero mean value, $S[\xi; \omega]$ is the Fourier transform of the correlation function:

$$S[\xi; \omega] = 2\int_{-\infty}^{\infty} e^{i\omega\tau} k(\tau)\, d\tau\,. \tag{2.10}$$

In the general case, where the mean value $\langle\xi\rangle = m$ is not zero, a similar formula holds after subtracting m from ξ:

$$S[\xi - m; \omega] \equiv 2\kappa(\omega) = 2\int_{-\infty}^{\infty} e^{i\omega\tau} k(\tau)\, d\tau\,. \tag{2.11}$$

Then, the spectral density $S[\xi; \omega]$ contains a delta function term at zero frequency. In fact, substituting $\langle\xi\xi_\tau\rangle = k(\tau) + m^2$ into (2.9), we obtain

$$S[\xi; \omega] = S[\xi - m; \omega] + 4\pi m^2\delta(\omega)\,. \tag{2.12}$$

It follows from (2.8) and (2.11) that the relation between the intensity coefficient and the spectral density is

$$K = \tfrac{1}{2} S[\xi - m; 0]\,. \tag{2.13}$$

If we take the inverse transform of (2.11), we find that

$$\begin{aligned} k(\tau) &= \frac{1}{4\pi}\int_{-\infty}^{\infty} e^{-i\omega\tau} S[\xi - m; \omega]\, d\omega \\ &= \frac{1}{2\pi}\int_0^{\infty} \cos\omega\tau\, S[\xi - m; \omega]\, d\omega\,. \end{aligned} \tag{2.14}$$

In particular, this gives the formula

$$\sigma^2(\xi) = k(0) = \frac{1}{2\pi}\int_0^{\infty} S[\xi - m; \omega]\, d\omega \tag{2.15}$$

for the variance of ξ. It can be shown that the spectral density is nonnegative for all ω, and hence can be interpreted physically as the distribution over frequency of the power of the process $\xi(t)$.

Inverting (2.9), we obtain

$$\langle \xi \xi_\tau \rangle = \frac{1}{4\pi} \int_{-\infty}^{\infty} \cos \omega\tau \, S[\xi; \omega] \, d\omega \,. \tag{2.16}$$

Sometimes it is convenient when calculating spectral densities to use techniques of operational calculus and the formula

$$S[\xi; \omega] = 4 \operatorname{Re} L[\langle \xi \xi_\tau \rangle; i\omega] \,, \tag{2.17}$$

which follows from (2.9); here, $L[f; p]$ denotes the Laplace transform

$$L[f; p] = \int_0^{\infty} e^{-pt} f(t) \, dt \,.$$

Example 1. Let $\xi(t)$ be an *exponentially correlated* random process, with correlation function

$$k(\tau) = \sigma^2 e^{-\beta|\tau|} \,. \tag{2.18}$$

Then, according to formulas (2.11) and (2.7), $\xi(t)$ has the spectral density

$$S[\xi - \langle \xi \rangle; \omega] = \frac{4\sigma^2\beta}{\omega^2 + \beta^2} \,, \tag{2.19}$$

and the correlation time

$$\tau_{cor} = \frac{1}{\beta} \,. \tag{2.20}$$

Example 2. If the correlation function has the form

$$k(\tau) = \sigma^2 e^{-\beta|\tau|} \cos \omega_0 \tau \,, \tag{2.21}$$

the spectral density equals

$$S[\xi - \langle \xi \rangle; \omega] = 4\beta\sigma^2 \frac{\omega^2 + \omega_0^2 + \beta^2}{(\omega^2 - \omega_0^2 - \beta^2)^2 + 4\beta^2\omega^2} \,. \tag{2.22}$$

If the inequality

$$\beta \ll \omega_0 \tag{2.23}$$

holds, the spectral density is concentrated in a relatively narrow

frequency band $\omega - \omega_0 \sim \beta$, and we have the approximations[1]

$$S[\xi - \langle\xi\rangle; \omega] \approx 4\beta\sigma^2 \frac{\omega^2 + \omega_0^2}{(\omega^2 - \omega_0^2)^2 + 4\beta^2\omega^2} \approx \frac{2\beta\sigma^2}{(\omega - \omega_0)^2 + \beta^2},$$
$$\tau_{cor} \approx \frac{2}{\pi} \cdot \frac{1}{\beta}. \tag{2.24}$$

The following table exhibits these and other commonly encountered correlation functions, together with the corresponding spectral densities:

TABLE I

No.	$k(\tau)$	$S[\xi - m; \omega] = 2\varkappa(\omega)$
1	$K\delta(\tau)$	$2K$
2	$\sigma^2 e^{-\beta\lvert\tau\rvert}$	$\dfrac{4\beta\sigma^2}{\omega^2 + \beta^2}$
3	$\sigma^2 e^{-1/4\beta^2\tau^2}$	$\sqrt{\pi}\dfrac{4\sigma^2}{\beta} e^{-\omega^2/\beta^2}$
4	$\sigma^2 e^{-\beta\lvert\tau\rvert} \cos \omega_0\tau$	$4\beta\sigma^2 \dfrac{\omega^2 + \omega_0^2 + \beta^2}{(\omega^2 - \omega_0^2 - \beta^2)^2 + 4\beta^2\omega^2}$
5	$e^{-\beta\lvert\tau\rvert}\left(\cos \omega_0\tau + \dfrac{\beta}{\omega_0} \sin \lvert \omega_0\tau \rvert\right)$	$4\beta \dfrac{\omega_0^2 + \beta^2}{(\omega^2 - \omega_0^2 - \beta^2)^2 + 4\beta^2\omega^2}$
6	$e^{-\lvert\beta\rvert\tau}\left(\cos \omega_0\tau - \dfrac{\beta}{\omega_0} \sin \lvert \omega_0\tau \rvert\right)$	$4\beta \dfrac{\omega^2}{(\omega^2 - \omega_0^2 - \beta^2)^2 + 4\beta^2\omega^2}$
7	$\sigma^2 \dfrac{\beta e^{-\alpha\lvert\tau\rvert} - \alpha e^{-\beta\lvert\tau\rvert}}{\beta - \alpha}$	$\dfrac{4\alpha\beta(\alpha + \beta)\sigma^2}{(\omega^2 + \alpha^2)(\omega^2 + \beta^2)}$
8	$\sigma^2 \dfrac{\sin \beta\tau}{\beta\tau}$	$\dfrac{2\pi\sigma^2}{\beta}$ if $\omega < \beta$; 0 if $\omega > \beta$

[1] The symbol $\approx$ means "is approximately equal to," while $\sim$ means "is of the order of."

2. Calculation of the Spectral Density by Using the Laplace Transform

Consider the Laplace transform of a stationary random process $\xi(t)$:

$$L[\xi; p] = \int_0^\infty e^{-pt}\, \xi(t)\, dt\,. \tag{2.25}$$

Multiplying this expression by its complex conjugate and averaging, we obtain

$$\langle |\, L[\xi; p]\, |^2 \rangle = \int_0^\infty \int_0^\infty e^{-pt}\, e^{-p^* t'}\, \langle \xi(t)\, \xi(t') \rangle\, dt\, dt'\,,$$

where the asterisk denotes the complex conjugate. Making the change of variables

$$t = s + \frac{\tau}{2}, \qquad t' = s - \frac{\tau}{2},$$

we find that

$$\langle |\, L[\xi; p]\, |^2 \rangle = \int_{-\infty}^{\infty} d\tau \int_{|\tau|/2}^{\infty} \exp\left\{ -(p + p^*)\, s - \frac{p - p^*}{2}\, \tau \right\} \langle \xi \xi_\tau \rangle\, ds\,.$$

Integrating with respect to s, we obtain

$$\langle |\, L[\xi; p]\, |^2 \rangle = \frac{1}{p + p^*} \int_{-\infty}^{\infty} \exp\left\{ -(p + p^*)\, \frac{|\tau|}{2} - \frac{p - p^*}{2}\, \tau \right\} \langle \xi \xi_\tau \rangle\, d\tau.$$

Multiplying by $p + p^*$ and comparing the result with (2.9), we arrive at the formula

$$S[\xi; \omega] = 2 \lim_{p \to -i\omega} (p + p^*)\, \langle |\, L[\xi; p]\, |^2 \rangle\,, \tag{2.26}$$

which is sometimes useful for calculating spectral densities.

Since the stationary process $\xi(t)$ does not grow at infinity, the Laplace transform (2.25) is defined for $\operatorname{Re} p > 0$, and so is the quantity $\langle\, |L[\xi; p]|^2 \rangle$. If we set

$$p = \frac{\epsilon}{2} - i\omega \qquad (\epsilon > 0)\,,$$

then the limit $p \to -i\omega$ corresponds to the limit $\epsilon \to 0$. Therefore, instead of (2.26), we can write

$$S[\xi; \omega] = 2 \lim_{\epsilon \to 0} \epsilon \left\langle \left| L\left[\xi; \frac{\epsilon}{2} - i\omega\right] \right|^2 \right\rangle . \tag{2.27}$$

3. The Random Spectrum and its Relation to the Spectral Density

The *spectrum* $y(\omega)$ of the process $\xi(t)$ is defined by the formula[2]

$$y(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega t}\, \xi(t)\, dt , \tag{2.28}$$

and is obviously a random function of the frequency ω.

★ Strictly speaking, when $\xi(t)$ is a stationary random process, the integral (2.28) diverges, so that (2.28) has only a formal meaning. In fact, (2.28) must then be interpreted as meaning that an ordinary equality is obtained if we apply a linear integral operator J_ω to both sides of (2.28), and then interchange the order of integration in the right-hand side:

$$J_\omega y(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (J_\omega e^{-i\omega t})\, \xi(t)\, dt . \tag{2.29}$$

Actually, this is the same way in which the familiar formal relation

$$\delta(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\nu\tau}\, d\nu \tag{2.30}$$

is interpreted. The expressions (2.28) and (2.30) are so-called *generalized functions*. Such a function is regarded as specified if we know the integral of the product of the function with an arbitrary continuous function. Moreover, it is not necessary that the value of a generalized function be specified for every value of its argument. As is well known, one can deal with generalized functions like (2.28) and (2.30) in the same way as with ordinary functions. ★

[2] The terms *spectrum* and *spectral density* are sometimes used as (rough) synonyms in the literature, but here their meanings are entirely different.

Multiplying (2.28) by its complex conjugate and averaging the result, we find that

$$\langle y(\omega)\, y^*(\omega')\rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i\omega t+i\omega' t'}\,\langle \xi(t)\,\xi(t')\rangle\, dt\, dt' . \qquad (2.31)$$

Making the change of variables

$$\tau = t' - t\,, \qquad t_0 = \frac{1}{2}(t + t')\,,$$

or equivalently

$$t = t_0 - \frac{\tau}{2}, \qquad t' = t_0 + \frac{\tau}{2},$$

and then integrating with respect to t_0, we obtain

$$\langle y(\omega)\, y^*(\omega')\rangle = \delta(\omega - \omega')\int_{-\infty}^{\infty} \exp\left\{i\,\frac{\omega + \omega'}{2}\,\tau\right\}\langle \xi\xi_\tau\rangle\, d\tau\,,$$

where we have used (2.30). According to (2.9), the remaining integral is just one half the spectral density $S[\xi;\omega]$, and hence

$$\langle y(\omega)\, y^*(\omega')\rangle = \frac{1}{2}\,\delta(\omega - \omega')\, S[\xi;\omega]\,. \qquad (2.32)$$

The inverse of the transformation (2.28) is

$$\xi(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i\omega t}\, y(\omega)\, d\omega\,. \qquad (2.33)$$

Since the process $\xi(t)$ is real, taking the complex conjugate of (2.28), we easily obtain

$$y^*(\omega) = y(-\omega)\,. \qquad (2.34)$$

We can then write (2.33) in the form

$$\xi(t) = \sqrt{\frac{2}{\pi}}\,\mathrm{Re}\int_0^{\infty} e^{i\omega t}\, y(\omega)\, d\omega\,, \qquad (2.35)$$

which is the customary way of writing oscillations in radio engineering (the method of complex amplitudes). Formula (2.35) represents the process $\xi(t)$ as a superposition of harmonic oscillations, where, in terms of the complex spectrum $y(\omega)$, the oscillation

of frequency ω has amplitude proportional to $|y(\omega)|$ and initial phase equal to $\arg y(\omega)$.

★ The random spectrum $y(\omega)$ appears only in the integrand of (2.35). Thus, to use this formula, relations of the type (2.29) are sufficient, and the strict equality (2.28) is not necessary. By the same token, to make our treatment more rigorous, we ought not to require that (2.31), (2.32) and (2.34) hold, but only that the corresponding formulas obtained by applying the integral operator J_ω hold. The use of integral operators is also justified physically, since experimentally we cannot obtain an exact harmonic, but can only isolate a sum of harmonics lying in a finite (albeit narrow) frequency band $\Delta\omega$.

Choosing J_ω to be the operation of integrating with respect to ω from ω_0 to $\omega_0 + \Delta\omega$, with weight $(1/\Delta\omega)e^{i\omega t_0}$, and writing

$$Y(\omega_0, t_0) = \frac{1}{\Delta\omega}\int_{\omega_0}^{\omega_0+\Delta\omega} e^{i\omega t_0}\, y(\omega)\, d\omega\,,$$

we obtain

$$Y(\omega_0, t_0) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \frac{1 - e^{-i\Delta\omega(t-t_0)}}{i\Delta\omega(t-t_0)}\, e^{-i\omega_0(t-t_0)}\, \xi(t)\, dt\,, \qquad (2.36)$$

according to (2.29). Carrying out similar integrations with respect to ω and ω' in both sides of (2.32), we obtain

$$\langle |Y(\omega_0, t_0)|^2\rangle = \frac{1}{2(\Delta\omega)^2}\int_{\omega_0}^{\omega_0+\Delta\omega} S[\xi;\omega]\, d\omega\,.$$

If $\Delta\omega$ is so small that the spectral density does not vary much in the interval from ω_0 to $\omega_0 + \Delta\omega$, this formula reduces to

$$\langle |Y(\omega_0, t_0)|^2\rangle = \frac{1}{2\Delta\omega}\, S[\xi;\omega_0]\,. \qquad (2.37)$$

The formula (2.37) shows the connection between the spectral density and the function $Y(\omega_0, t_0)$, which to a certain approximation replaces the true spectrum $y(\omega_0)e^{-i\omega_0 t_0}$. If the bandwidth $\Delta\omega$ is small, the factor

$$\frac{1 - e^{-i\Delta\omega(t-t_0)}}{i\Delta\omega(t-t_0)}$$

in (2.36) is slowly varying. This factor has the effect of restricting the region of integration in the divergent integral (2.28), and as a result, the influence of values of t for which $|t - t_0| \gg 1/\Delta\omega$ becomes negligibly small. Of course, there are many other ways of restricting the region of integration in (2.28), corresponding to other forms of the integral operator J_ω in (2.29).★

Under actual conditions, one does not deal with the exact spectrum $y(\omega)$. Thus, consider the function

$$Y(\omega_0, t_0) = \int_{-\infty}^{\infty} G(t_0 - t)\, \xi(t)\, dt\,, \tag{2.38}$$

which is a good approximation to the spectrum $y(\omega)$ if the function $G(t_0 - t)$ is "narrow-band." We shall regard the function $G(t_0 - t)$ as being narrow-band if in the spectral expansion

$$G(t_0 - t) = \frac{1}{\sqrt{2\pi}} \int e^{i\omega(t_0 - t)}\, g(\omega)\, d\omega \qquad \left(\int g(\omega)\, d\omega = 1\right), \tag{2.39}$$

the function $g(\omega)$ is appreciably different from zero only in a small interval

$$|\omega - \omega_0| \sim \Delta\omega \qquad (\Delta\omega \ll \omega_0) \tag{2.40}$$

centered at ω_0. This means that $G(t_0 - t)$ is just the harmonic $e^{i\omega_0(t_0 - t)}$ multiplied by a slowly varying function, which changes in a time of order $1/\Delta\omega$. Substituting (2.39) into (2.38), changing the order of integration and taking account of (2.28), we have

$$Y(\omega_0, t_0) = \int e^{i\omega t_0}\, g(\omega)\, y(\omega)\, d\omega\,.$$

Multiplying (2.32) by $g(\omega)g^*(\omega')e^{i(\omega - \omega')t}$ and integrating, we obtain

$$\langle |\, Y(\omega_0, t_0)\,|^2 \rangle = \tfrac{1}{2} \int S[\xi;\, \omega]\, |\, g(\omega)\,|^2\, d\omega\,.$$

Suppose that the function $g(\omega)$ is appreciably different from zero only in an interval (2.40) which is so narrow that $S[\xi;\, \omega]$ practically

does not change there. Then, we can bring the function $S[\xi; \omega]$ in front of the integral sign, obtaining

$$\langle | Y(\omega_0, t_0) |^2 \rangle = \tfrac{1}{2} S[\xi; \omega_0] \int | g(\omega) |^2 \, d\omega \,. \tag{2.41}$$

According to formula (2.41), to obtain the spectral density it is necessary (in radio engineering language) to pass the random function $\xi(t)$ through a narrow band linear filter followed by a quadratic detector, and then average the result. However, in this regard, it should be pointed out that any actual radio engineering device has a real impulse response $G(t_0 - t)$. This means that $g(\omega)$ has the property

$$g(-\omega) = g^*(\omega) \,,$$

so that $g(\omega)$ must be different from zero not only in the interval (2.40) but also in the symmetric interval $| \omega + \omega_0 | \sim \Delta\omega$, where, because of its evenness, the spectral density $S[\xi; \omega]$ takes the same values as in (2.40). Therefore, as before, we can bring $S[\xi; \omega]$ in front of the integral sign. Thus, in the case of a real narrow-band process $Y(\omega_0, t_0)$, formula (2.41) remains the same.

4. Time Averages of Stationary Processes

The so-called *ergodic processes* are an important subclass of the class of stationary processes. If we know a single realization of an ergodic process over the whole time axis $-\infty < t < \infty$, then by making time shifts, we can obtain an infinite statistical ensemble of realizations. Therefore, from one realization of an ergodic process, we can ascertain all of its statistical characteristics.

For example, the mean value of $\xi(t)$ can be calculated by the formula

$$m = \lim_{T \to \infty} \frac{1}{T} \int_0^T \xi(t) \, dt \,. \tag{2.42}$$

Of course, other formulas can also be used, in particular, the formula

$$m = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{\infty} \xi(t + k\Delta) , \tag{2.43}$$

which replaces the integral (2.42) by a sum, where Δ is any positive number. To prove (2.42), consider the random variable

$$m_T = \frac{1}{T} \int_0^T \xi(t)\, dt , \tag{2.44}$$

which has mean value $\langle m_T \rangle = m$. As we shall now show, the variance $\mathbf{D}m_T$ goes to zero as $T \to \infty$.

Squaring (2.44), averaging the result and subtracting m^2, we obtain

$$\mathbf{D}m_T = \frac{1}{T^2} \int_0^T \int_0^T k(t - t')\, dt\, dt' .$$

Making the change of variables

$$\tau = t - t' , \qquad t_0 = \tfrac{1}{2}(t + t') ,$$

we see that for a fixed value of τ, the limits of integration with respect to t_0 go from $\frac{1}{2}|\tau|$ to $T - \frac{1}{2}|\tau|$. Thus, integrating first with respect to t_0 and then with respect to τ, we obtain

$$\mathbf{D}m_T = \frac{1}{T^2} \int_{-T}^{T} (T - |\tau|)\, k(\tau)\, d\tau , \tag{2.45}$$

or because of (2.5),

$$\mathbf{D}m_T = \frac{2\sigma^2}{T} \int_0^T \left(1 - \frac{\tau}{T}\right) R(\tau)\, d\tau .$$

Since $\mathbf{D}m_T$ is nonnegative, it follows that

$$\mathbf{D}m_T = \frac{2\sigma^2}{T} \left| \int_0^T \left(1 - \frac{\tau}{T}\right) R(\tau)\, d\tau \right| \leqslant \frac{2\sigma^2}{T} \int_0^T \left(1 - \frac{\tau}{T}\right) |R(\tau)|\, d\tau$$

$$\leqslant \frac{2\sigma^2}{T} \int_0^T |R(\tau)|\, d\tau \leqslant \frac{2\sigma^2}{T} \int_0^\infty |R(\tau)|\, d\tau .$$

The last integral is the correlation time (2.7), and hence

$$\mathrm{D}m_T \leqslant 2\sigma^2 \frac{\tau_{cor}}{T}. \tag{2.46}$$

From (2.46) it is clear that if the stationary process $\xi(t)$ has a finite correlation time and a finite variance $\sigma^2 = \langle \xi^2(t) \rangle - m^2$, then

$$\mathrm{D}m_T \to 0 \quad \text{as} \quad T \to \infty.$$

This means that as T increases, the random variable m_T converges to a quantity which is not random, i.e., the mean value m. The formula (2.46) also leads to the following rough estimate of the speed of convergence:

$$\left| \frac{1}{T} \int_0^T \xi(t)\, dt - m \right| \lesssim \sigma \left(2 \frac{\tau_{cor}}{T} \right)^{1/2}. \tag{2.47}$$

The time average (2.44) converges at the same rate as the arithmetic mean of a family of N identically distributed, independent random variables $\xi(t_i)$, where $N = T/2\tau_{cor}$. In fact, in order to simplify practical calculations of the mean value, it is convenient to go from the integral (2.42) to the sum (2.43) by choosing $\Delta \approx 2\tau_{cor}$.

Just as in the case of the mean value, we can use time averaging to calculate other statistical characteristics of ergodic processes. For example, forming the new stationary random function

$$\eta(t) = \xi(t)\, \xi(t + \tau) \tag{2.48}$$

(where τ is fixed), and applying formula (2.42) to $\eta(t)$, we find that

$$\langle \xi \xi_\tau \rangle = \lim_{T \to \infty} \frac{1}{T} \int_0^T \xi(t)\, \xi(t + \tau)\, dt, \tag{2.49}$$

a result which can be used when $\eta(t)$ has a finite correlation time.

If we want to calculate the spectral density, then, according to (2.41), we have to time-average the square of the function $Y(\omega_0, t_0)$. Both the process $Y(\omega_0, t_0)$ and its square are stationary processes, regarded as functions of the time t_0. Since $Y(\omega_0, t_0)$ is more narrow-band than the original process $\xi(t)$, it has a longer correlation time,

whose order of magnitude is inversely proportional to the bandwidth $\Delta\omega$. It is natural to assume that $|\, Y(\omega_0, t_0)\,|^2$ has approximately the same correlation time:

$$(\tau_{c\omega})_{Y^2} \sim 1/\Delta\omega\,.$$

Applying formula (2.47) to the process $|\, Y(\omega_0, t_0)\,|^2$, we have

$$\left| \langle |\, Y(\omega_0, t_0)\,|^2 \rangle - \frac{1}{T}\int_0^T |\, Y(\omega_0, t_0)\,|^2\, dt_0 \right| \lesssim \left(\frac{2D\,|\, Y\,|^2}{T\Delta\omega}\right)^{1/2}. \qquad (2.50)$$

or taking account of (2.41),

$$\left| S[\xi; \omega_0] - \frac{2}{\int |\, g\,|^2\, d\omega}\,\frac{1}{T}\int_0^T |\, Y(\omega_0, t_0)\,|^2\, dt_0 \right| \sim \frac{\sigma(|\, Y\,|^2)}{\int |\, g\,|^2\, d\omega}\,\frac{1}{\sqrt{T\Delta\omega}}\,.$$

It can be shown that the estimate

$$\frac{\sigma(|\, Y\,|^2)}{\int |\, g\,|^2\, d\omega} \sim S[\xi; \omega_0]$$

holds. This implies that

$$S[\xi; \omega_0] = \frac{2}{\int |\, g\,|^2\, d\omega}\,\frac{1}{T}\int_0^T |\, Y(\omega_0, t)\,|^2\, dt\,, \qquad (2.51)$$

with a relative error of order $(T\Delta\omega)^{-1/2}$. We obtain the exact value of the spectral density if we pass to the limits $T\Delta\omega \to \infty$ and $\Delta\omega \to 0$ in (2.51).

In making experimental measurements of the spectral density, it is not desirable to use a very small bandwidth $\Delta\omega$, since to obtain satisfactory accuracy the averaging time T has to be very much greater than the time $1/\Delta\omega$ required for the narrow-band process to change appreciably. Of course, the operation of time averaging need not take the explicit form $(1/T)\int_0^T \ldots\, dt$. In fact, as a rule, the integrators which perform the time averaging in practice have other weight functions. The important thing is that the integration time T should be large enough to satisfy the inequalities

$$T \gg \tau_{\mathrm{corr}}\,, \qquad T \gg \frac{1}{\Delta\omega}\,.$$

5. Cross-Correlation Functions

Suppose we are given several random functions $\xi(t)$, $\eta(t)$, ..., which for full generality, we assume to be complex. If the processes are stationary, then their correlation functions

$$\mathbf{K}[\xi(t_1), \xi^*(t_2)], \qquad \mathbf{K}[\eta(t_1), \eta^*(t_2)] \tag{2.52}$$

depend only on the time difference $t_2 - t_1$. In addition to the correlation functions (2.52), we are interested in cross correlations between different random functions, e.g., $\xi(t)$ and $\eta(t)$. Two stationary processes $\xi(t)$ and $\eta(t)$ are said to be *stationarily correlated* if the function

$$\mathbf{K}[\xi(t_1), \eta^*(t_2)] = k_{\xi\eta}(t_2 - t_1) \tag{2.53}$$

also depends only on the time difference $t_2 - t_1$. The function (2.53) is called the *cross-correlation function*, as opposed to the functions (2.52), which are sometimes called *autocorrelation functions*. Unlike an autocorrelation function, a cross-correlation function need not be an even function of its argument $t_2 - t_1$.

If the mean values $\langle\xi\rangle$ and $\langle\eta\rangle$ equal zero, then the Fourier transform of (2.53) gives the *cross-spectral density*

$$S[\xi, \eta; \omega] = 2\int_{-\infty}^{\infty} e^{i\omega\tau}\langle\xi\eta_\tau^*\rangle\, d\tau\,, \tag{2.54}$$

which is not necessarily real, but has the property

$$S^*[\xi, \eta; \omega] = S[\eta, \xi; \omega]\,. \tag{2.55}$$

Moreover, for real processes, we have the relation

$$S^*[\xi, \eta; \omega] = S[\xi, \eta; -\omega]\,. \tag{2.56}$$

If we set $\eta \equiv \xi$ in (2.54), we obtain the spectral density of a single random function. In the notation of (2.54), the spectral density $S[\xi; \omega]$ defined by formula (2.9) can be written as $S[\xi, \xi; \omega]$. The inverse of the transformation (2.54) is

$$\langle\xi\eta_\tau^*\rangle = \frac{1}{4\pi}\int_{-\infty}^{\infty} S[\xi, \eta; \omega]\, e^{-i\omega\tau}\, d\omega\,. \tag{2.57}$$

We now consider the random spectra

$$y_\xi(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega t}\, \xi(t)\, dt\,,$$
$$y_\eta(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega t}\, \eta(t)\, dt\,, \tag{2.58}$$

in terms of which we can write the processes $\xi(t)$ and $\eta(t)$ as superpositions of harmonic oscillations, by using expressions like (2.33), or like (2.35) in the case of real processes. By the same argument as used to derive formula (2.32), we can easily obtain the formula

$$\langle y_\xi(\omega)\, y_\eta^*(\omega')\rangle = \tfrac{1}{2}\, \delta(\omega - \omega')\, S[\xi, \eta; \omega] \tag{2.59}$$

from (2.58) and (2.54). Next, we introduce the narrow-band random processes

$$Y_\xi(\omega_0, t_0) = \int e^{i\omega t_0}\, g_\xi(\omega)\, y_\xi(\omega)\, d\omega = \int G_\xi(t_0 - t)\, \xi(t)\, dt\,,$$
$$Y_\eta(\omega_0, t_0) = \int e^{i\omega t_0}\, g_\eta(\omega)\, y_\eta(\omega)\, d\omega = \int G_\eta(t_0 - t)\, \eta(t)\, dt\,, \tag{2.60}$$

which, like (2.38), replace the exact spectra. Here, the functions g_ξ, G_ξ, g_η, G_η have the same meaning as before, but in the general case, they may not be the same for the processes ξ and η. Using (2.60) and (2.59), we easily find that

$$\langle Y_\xi(\omega_0, t_0)\, Y_\eta^*(\omega_0, t_0)\rangle = \tfrac{1}{2} \int S[\xi, \eta; \omega]\, g_\xi(\omega)\, g_\eta^*(\omega)\, d\omega\,. \tag{2.61}$$

The presence of the functions g_ξ, g_η in (2.60) converts ξ, η into the narrow-band signals Y_ξ, Y_η. If g_ξ and g_η have the form

$$g_\xi(\omega) = g_\eta(\omega) = g(\omega)\,, \tag{2.62}$$

where $g(\omega)$ is a real function which is appreciably different from zero only in a narrow frequency band $\omega - \omega_0 \sim \Delta\omega$, then

$$\langle Y_\xi(\omega_0, t_0)\, Y_\eta^*(\omega_0, t_0)\rangle = \tfrac{1}{2}\, S[\xi, \eta; \omega_0] \int g^2(\omega)\, d\omega\,. \tag{2.63}$$

Therefore, to obtain the (complex) value of the cross-spectral density, we have to take the average of the complex quantity

$$Y_\xi Y_\eta^* = | Y_\xi | | Y_\eta | e^{i\Delta\varphi} . \tag{2.64}$$

Here, we have to take into account not only the product $| Y_\xi | | Y_\eta |$ of the amplitudes of the narrow-band signals, but also their phase difference $\Delta\varphi$, which determines the argument of the complex number (2.64). The processes $\xi(t)$, $\eta(t)$ dealt with in radio engineering are real, and so are the transformations converting them into real narrow-band processes Y_ξ, Y_η. This imposes the restrictions

$$g_\xi(-\omega) = g_\xi^*(\omega) , \qquad g_\eta(-\omega) = g_\eta^*(\omega) \tag{2.65}$$

on g_ξ and g_η. If we set

$$g_\xi(\omega) = g_\eta(\omega) = g(| \omega |) ,$$

where g is a real function, formula (2.61) gives

$$\langle Y_\xi(\omega_0, t_0)\, Y_\eta(\omega_0, t_0)\rangle = \operatorname{Re} S[\xi, \eta; \omega_0] \int_0^\infty g^2(\omega)\, d\omega . \tag{2.66}$$

We see that forming the product of two real, narrow-band processes cannot give complete information about the cross-spectral density as a complex function. To complete the specification of $S[\xi, \eta; \omega_0]$, we need a second pair of narrow-band processes, but only one function in the first pair need be changed. Writing

$$g_\xi(\omega) = g(| \omega |), \qquad \tilde{g}_\eta(\omega) = -ig\,(| \omega |) \operatorname{sgn} \omega ,$$

we find from (2.61) that

$$\langle Y_\xi(\omega_0, t_0)\, \tilde{Y}_\eta(\omega_0, t_0)\rangle = \operatorname{Im} S[\xi, \eta; \omega_0] \int_0^\infty g^2(\omega)\, d\omega . \tag{2.67}$$

Thus, to experimentally determine the cross-spectral density, it is not enough to form just one product of two narrow-band processes. Another product of the two processes must be formed, and before doing so, we have to use a phase inverter to introduce an extra phase shift (preferably of 90°) between the two processes.

CHAPTER 3

Gaussian and Non-Gaussian Random Processes. Quasi-Moment Functions

1. Gaussian Processes and Their Probability Densities

Suppose that all the correlation functions $k_2(t_1, t_2)$, $k_3(t_1, t_2, t_3)$, ... of a random process $\xi(t)$ are zero, except the first-order correlation function $k_1(t_1)$. Then, $\xi(t)$ is not random at all, and is said to be *deterministic*, which means that all the realizations of $\xi(t)$ are the same. Next, suppose that all the higher-order correlation functions of $\xi(t)$ are zero, except the first-order and second-order correlation functions $k_1(t_1)$ and $k_2(t_1, t_2)$. Then, the process $\xi(t)$ is said to be *Gaussian* (or *normal*). In this case, according to formula (1.67), we have the following expression for the characteristic function of the n random variables $\xi(t_1), \ldots, \xi(t_n)$:

$$\Theta_n(u_1, \ldots, u_n; t_1, \ldots, t_n) = \exp\left\{i\sum_{\alpha=1}^{n} k_1(t_\alpha)\, u_\alpha - \tfrac{1}{2}\sum_{\alpha,\beta=1}^{n} k_2)t_\alpha, t_\beta)\, u_\alpha u_\beta\right\}. \quad (3.1)$$

Using (3.1), we can find the probability density

$$w_n[\xi(t_1), \ldots, \xi(t_n)]$$

by calculating the inverse transformation

$$w_n(\xi_1, \ldots, \xi_n) \quad (3.2)$$
$$= \frac{1}{(2\pi)^n}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp\left\{-i\sum_{\alpha=1}^{n} u_\alpha\xi_\alpha\right\}\Theta_n(u_1, \ldots, u_n; t_1, \ldots, t_n)\, du_1 \ldots du_n ,$$

where $\xi_\alpha = \xi(t_\alpha)$; this is just the multidimensional generalization of the integral (1.19). The expression $\Sigma_{\alpha,\beta}\, k_2(t_\alpha, t_\beta)\, u_\alpha u_\beta$, which can be written as

$$\left\langle \left\{ \sum_{\alpha=1}^{n} [\xi(t_\alpha) - k_1(t_\alpha)]\, u_\alpha \right\}^2 \right\rangle,$$

is a positive definite quadratic form. Therefore, the absolute value of the characteristic function appearing in (3.2) certainly does not increase as $u_1, \ldots, u_n$ get larger. This implies that regardless of the presence of the infinite limits of integration, the integral (3.2) exists, at least in the same sense as the integral (2.30). Actually, we can carry out the integration in (3.2) explicitly, obtaining

$$\begin{aligned} w_n(\xi_1, \ldots, \xi_n) &= (2\pi)^{-n/2} [\det \| k_2(t_\alpha, t_\beta) \|]^{-1/2} \\ &\times \exp\left\{ -\tfrac{1}{2} \sum_{\alpha,\beta=1}^{n} a_{\alpha\beta} [\xi_\alpha - k_1(t_\alpha)]\, [\xi_\beta - k_1(t_\beta)] \right\}, \end{aligned} \tag{3.3}$$

where $\det \| k_2(t_\alpha, t_\beta) \|$ is the determinant of the *correlation matrix*

$$\| k_2(t_\alpha, t_\beta) \| = \begin{Vmatrix} k_2(t_1, t_1) & k_2(t_1, t_2) & \ldots & k_2(t_1, t_n) \\ \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \ldots & \cdot \\ k_2(t_n, t_1) & k_2(t_n, t_2) & \ldots & k_2(t_n, t_n) \end{Vmatrix}, \tag{3.4}$$

and $\| a_{\alpha\beta} \|$ is the inverse of the correlation matrix, i.e.,

$$\| a_{\alpha\beta} \| = \| k_2(t_\alpha, t_\beta) \|^{-1}, \qquad \sum_{\beta=1}^{n} a_{\alpha\beta} k_2(t_\beta, t_\gamma) = \delta_{\alpha\gamma} \tag{3.5}$$

($\delta_{\alpha\gamma} = 1$ for $\alpha = \gamma$, $\delta_{\alpha\gamma} = 0$ if $\alpha \neq \gamma$). A probability density of the form (3.3) is said to be *Gaussian* (or *normal*), and the random variables $\xi_1, \ldots, \xi_n$ are said to have a *Gaussian* (or *normal*) distribution.

If $\xi(t)$ is stationary, and if we set $n = 1$, the matrix (3.4) and its inverse each reduce to a single element, and in fact

$$a_{11} = \frac{1}{k_2(t_1, t_1)} = \frac{1}{\sigma^2}$$

[cf. (2.5)]. Therefore, the one-dimensional probability density equals

$$w[\xi(t)] = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2\sigma^2}[\xi(t) - m]^2\right\}. \tag{3.6}$$

Setting $n = 2$, we find that the inverse of the correlation matrix is

$$\left\| \begin{matrix} k(0) & k(\tau) \\ k(\tau) & k(0) \end{matrix} \right\|^{-1} = \left\| \begin{matrix} \dfrac{k(0)}{k^2(0) - k^2(\tau)} & \dfrac{-k(\tau)}{k^2(0) - k^2(\tau)} \\ \dfrac{-k(\tau)}{k^2(0) - k^2(\tau)} & \dfrac{k(0)}{k^2(0) - k^2(\tau)} \end{matrix} \right\|, \tag{3.7}$$

so that the two-dimensional probability density is

$$w_2(\xi_1, \xi_2) = \frac{1}{2\pi\sigma^2\sqrt{1 - R^2(\tau)}} \exp\left\{-\frac{1}{2\sigma^2}\frac{\xi_1^2 - 2R(\tau)\,\xi_1\xi_2 + \xi_2^2}{1 - R^2(\tau)}\right\}. \tag{3.8}$$

In certain problems where it is difficult to calculate integrals involving the probability density (3.8), it is convenient to expand (3.8) in a series of powers of the correlation coefficient $R(\tau)$. To find such an expansion, we set $n = 2$, $k_1 = m = 0$ in formulas (3.2) and (3.1), obtaining

$$w_2(\xi_1, \xi_2) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left\{-iu_1\xi_1 - iu_2\xi_2 - \frac{\sigma^2}{2}[u_1^2 + 2R(\tau)\,u_1u_2 + u_2^2]\right\} du_1\,du_2. \tag{3.9}$$

Using the expansion

$$e^{-\sigma^2 R u_1 u_2} = \sum_{p=0}^{\infty} \frac{(-1)^p}{p!}(\sigma^2 R u_1 u_2)^p, \tag{3.10}$$

we find that

$$w_2(\xi_1, \xi_3) = \sum_{p=0}^{\infty}(-1)^p\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} u_1^p \exp\left\{-iu_1\xi_1 - \frac{\sigma^2u_1^2}{2}\right\} du_1\right] \times \left[\frac{1}{2\pi}\int_{-\infty}^{\infty} u_2^p \exp\left\{-iu_2\xi_2 - \frac{\sigma^2u_2^2}{2}\right\} du_2\right]\frac{\sigma^{2p}R^p}{p!}. \tag{3.11}$$

Making the change of variables

$$\lambda_1 = \sigma u_1 , \qquad \lambda_2 = \sigma u_2,$$

and using the integral representation

$$F^{(p+1)}(z) = \frac{(-i)^p}{2\pi} \int_{-\infty}^{\infty} \lambda^p \exp\left\{-i\lambda z - \frac{\lambda^2}{2}\right\} d\lambda \tag{3.12}$$

for the derivatives

$$F^{(p+1)}(z) = \frac{d^{p+1}F(z)}{dz^{p+1}} = \frac{1}{\sqrt{2\pi}} \frac{d^p}{dz^p} e^{-z^2/2} \tag{3.13}$$

of the *probability integral*

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2}\, dx ,$$

we obtain the expansion[1]

$$w(\xi_1, \xi_2) = \frac{1}{\sigma^2} \sum_{p=0}^{\infty} F^{(p+1)}\left(\frac{\xi_1}{\sigma}\right) F^{(p+1)}\left(\frac{\xi_2}{\sigma}\right) \frac{R^p}{p!} . \tag{3.14}$$

Formula (3.14) can also be written in terms of Hermite polynomials.

Similar expansions exist for the higher-order normal distributions, i.e., for the normal distributions of order three, four, etc. For example, making series expansions like (3.10) of the exponentials

$$e^{-\sigma^2 R_{12} u_1 u_2}, \qquad e^{-\sigma^2 R_{23} u_2 u_3}, \qquad e^{-\sigma^2 R_{13} u_1 u_3},$$

where

$$R_{12} = R(t_1 - t_2), \qquad R_{23} = R(t_2 - t_3), \qquad R_{13} = R(t_1 - t_3),$$

[1] See H. Cramér, *Methods of Mathematical Statistics*, Princeton University Press, Princeton, N.J. (1946), p. 290, where the result in question is attributed to early papers by Stieltjes and Sheppard.

we can write the three-dimensional characteristic function in the form

$$\Theta_3(u_1, u_2, u_3) = \exp\left\{-\frac{\sigma^2}{2}(u_1^2 + u_2^2 + u_3^2)\right\} \sum_{p=0}^{\infty} (-1)^p \frac{\sigma^{2p}}{p!} (R_{12}u_1u_2)^p \times \sum_{q=0}^{\infty} (-1)^q \frac{\sigma^{2q}}{q!} (R_{23}u_2u_3)^q \sum_{r=0}^{\infty} (-1)^r \frac{\sigma^{2r}}{r!} (R_{13}u_1u_3)^r . \tag{3.15}$$

In calculating the Fourier transform of (3.15), each term with fixed values of p, q, r reduces to a product of three integrals of the type (3.12). Therefore, we find that

$$w(\xi_1, \xi_2, \xi_3) = \frac{1}{\sigma^3} \sum_{p,q,r=0}^{\infty} F^{(p+r+1)}\left(\frac{\xi_1}{\sigma}\right) F^{(p+q+1)}\left(\frac{\xi_2}{\sigma}\right) F^{(q+r+1)}\left(\frac{\xi_3}{\sigma}\right) \frac{R_{12}^p R_{23}^q R_{13}^r}{p!\, q!\, r!} . \tag{3.16}$$

According to what has been said, the most important (and the most time-consuming) step in finding a multidimensional probability density for a normal process consists in finding the matrix $\| a_{\alpha\beta} \|$ which is the inverse of the correlation matrix $\| k_2(t_\alpha, t_\beta) \|$. The problem of calculating the inverse matrix is equivalent to the problem of solving the system of equations

$$\begin{aligned} k_2(t_1, t_1)\, y_1 + k_2(t_1, t_2)\, y_2 + \ldots + k_2(t_1, t_n)\, y_n &= x_1 , \\ \ldots \quad \ldots \quad \ldots \quad \ldots \quad &\ldots \\ k_2(t_n, t_1)\, y_1 + k_2(t_n, t_2)\, y_2 + \ldots + k_2(t_n, t_n)\, y_n &= x_n . \end{aligned} \tag{3.17}$$

If we solve (3.17) with respect to $y_1, \ldots, y_n$, we obtain expressions of the form

$$y_\alpha = \sum_{\beta=1}^{n} a_{\alpha\beta} x_\beta \qquad (\alpha = 1, \ldots, n) . \tag{3.18}$$

Using (3.18), we can easily find the quadratic form

$$\sum_{\alpha,\beta=1}^{n} a_{\alpha\beta} [\xi_\alpha - k_1(t_\alpha)]\, [\xi_\beta - k_1(t_\beta)]$$

appearing in (3.3). In fact, we need only form the sum

$$2\varphi[x] = \sum_{\alpha=1}^{n} x_\alpha y_\alpha = \sum_{\alpha,\beta=1}^{n} a_\alpha x_{\beta\alpha} x_\beta , \tag{3.19}$$

and then set $x_\alpha = \xi_\alpha - k_1(t_\alpha)$.

This method is applicable for any number of instants of time $t_1, \ldots, t_n$, and remains meaningful even when the points $t_1, \ldots, t_n$ lie arbitrarily close together. In this case, the finite probability density $w_n[\xi(t_1), \ldots, \xi(t_n)]$ goes over into a *probability functional.*

2. Conditional Correlation Functions of Gaussian Processes

Suppose we know the values $\xi(T_1), \ldots, \xi(T_s)$ of a random process $\xi(t)$ at certain time instants $T_1, \ldots, T_s$. This information changes the initial data concerning the random process, in particular, its correlation functions and probability densities. The new, modified correlation functions and probability densities are said to be *conditional* (or *a posteriori*), as opposed to the original quantities, which are said to be *unconditional* (or *a priori*). Qualitatively speaking, as a result of knowing certain values of a random process, we obtain a new "a posteriori random process," which is characterized by different statistical properties. It should be noted that if the a priori random process is normal, then so is the a posteriori random process.

According to formula (1.39), the conditional probability density of the random variables $\xi(t_1), \ldots, \xi(t_r)$, given the values $\xi(T_1), \ldots, \xi(T_s)$, is equal to

$$w[\xi(t_1), \ldots, \xi(t_r) \mid \xi(T_1), \ldots, \xi(T_s)] = \frac{1}{N} w_{r+s}[\xi(t_1), \ldots, \xi(t_r), \xi(T_1), \ldots, \xi(T_s)] , \tag{3.20}$$

where $1/N$ is a normalization constant, which does not depend on

$\xi(t_1), \ldots, \xi(t_r)$. The Fourier transform of (3.20) with respect to the variables $\xi(t_1), \ldots, \xi(t_r)$ gives the *conditional characteristic function*

$$\Theta_{cond}(u_1, \ldots, u_r) = \frac{1}{N} \int \ldots \int \exp\left\{ i \sum_{\alpha=1}^{r} u_\alpha \xi(t_\alpha) \right\}$$

$$\times\, w_{r+s}[\xi(t_1), \ldots, \xi(t_r), \xi(T_1), \ldots, \xi(T_s)]\, d\xi(t_1) \ldots d\xi(t_r).$$

Expressing the last integral in terms of the a priori characteristic function $\Theta_{r+s}(u_1, \ldots, u_r, v_1, \ldots, v_s)$, which is the $(r+s)$-dimensional Fourier transform of the joint probability density

$$w_{r+s}[\xi(t_1), \ldots, \xi(t_r), \xi(T_1), \ldots, \xi(T_s)]\,,$$

we find that

$$\Theta_{cond}(u_1, \ldots, u_r) = \frac{1}{N(2\pi)^s} \int \ldots \int \exp\left\{ -i \sum_{\beta=1}^{s} v_\beta \xi(T_\beta) \right\} \tag{3.21}$$

$$\times\, \Theta_{r+s}(u_1, \ldots, u_r, v_1, \ldots, v_s)\, dv_1 \ldots dv_s\,.$$

The factor $1/N$ in (3.21) is determined by the condition $\Theta_{cond}(0, \ldots, 0) = 1$.

The calculation of the integral (3.21) is completely analogous to the calculation of the integral (3.2). The only difference is that there is now extra dependence on the numbers $u_1, \ldots, u_r$, which can be regarded as parameters. According to (3.1), the characteristic function appearing in (3.21) can be written in the form

$$\Theta_{r+s}(u_1, \ldots, u_r, v_1, \ldots, v_s) = \exp\left\{ i \sum_{\alpha=1}^{r} k_1(t_\alpha)\, u_\alpha + i \sum_{\beta=1}^{s} k_1(T_\beta)\, v_\beta \right.$$

$$\left. - \tfrac{1}{2} \sum_{\alpha,\beta=1}^{r} k_2(t_\alpha, t_\beta)\, u_\alpha u_\beta + i^2 \sum_{\alpha=1}^{r} \sum_{\beta=1}^{s} k_2(t_\alpha, T_\beta)\, u_\alpha v_\beta - \tfrac{1}{2} \sum_{\alpha,\beta=1}^{s} k_2(T_\alpha, T_\beta)\, v_\alpha v_\beta \right\}. \tag{3.22}$$

If we let $\| a_{\alpha\beta} \|$ denote the matrix which is the inverse of the

correlation matrix $\| k_2(T_\alpha, T_\beta) \|$, then it follows from (3.21), by analogy with (3.3), that

$$\Theta_{cond}(u_1, \ldots, u_r) = \frac{1}{N'} \exp \left\{ i \sum_{\alpha=1}^{r} k_1(t_\alpha) u_\alpha - \tfrac{1}{2} \sum_{\alpha,\beta=1}^{r} k_2(t_\alpha, t_\beta) u_\alpha u_\beta \right.$$

$$- \tfrac{1}{2} \sum_{\alpha,\beta=1}^{s} a_{\alpha\beta} \Big[\xi(T_\alpha) - k_1(T_\alpha) - i \sum_{\gamma=1}^{r} k_2(T_\alpha, t_\gamma) u_\gamma \Big] \qquad (3.23)$$

$$\left. \times \Big[\xi(T_\beta) - k_1(T_\beta) - i \sum_{\delta=1}^{r} k_2(T_\beta, t_\delta) u_\delta \Big] \right\}.$$

The factor

$$(2\pi)^{-r/2} [\det \| k_2(T_\alpha, T_\beta) \|]^{-1/2}$$

has been incorporated into the factor $1/N'$, which is independent of $u_1, \ldots, u_r$.

We now introduce the *conditional correlation functions* $\tilde{k}_1(t)$ and $\tilde{k}_2(t, t')$, in terms of which the conditional characteristic function (3.23) can be written in the same form as (3.1), i.e.,

$$\Theta_{cond}(u_1, \ldots, u_r) = \exp \left\{ i \sum_{\alpha=1}^{r} \tilde{k}_1(t_\alpha) u_\alpha - \tfrac{1}{2} \sum_{\alpha,\beta=1}^{r} \tilde{k}_2(t_\alpha, t_\beta) u_\alpha u_\beta \right\}. \qquad (3.24)$$

Separately equating the coefficients of the linear and quadratic forms in (3.23) and (3.24), we obtain explicit expressions for $k_1(t)$ and $k_2(t, t')$:

$$\tilde{k}_1(t) = k_1(t) + \sum_{\alpha,\beta=1}^{s} k_2(t, T_\alpha) a_{\alpha\beta} [\xi(T_\beta) - k_1(T_\beta)] ,$$

$$\tilde{k}_2(t, t') = k_2(t, t') - \sum_{\alpha,\beta=1}^{s} k_2(t, T_\alpha) a_{\alpha\beta} k_2(T_\beta, t') . \qquad (3.25)$$

These expressions completely determine the a posteriori random process, just as in the case of any other Gaussian process.[2]

[2] Formula (3.25) can also be applied (with an obvious modification) to the case where we know a continuous realization of the random process on an entire interval.

As an example, consider a stationary process with $k_1(t) = m$, $k_2(t, t') = \sigma^2 R(t - t')$, and suppose we know the value of the process at some time t_0. Then, according to (3.25), we have the following conditional mean value and conditional second-order correlation function:

$$\tilde{k}_1(t) = m + R(t - t_0)\,[\xi(t_0) - m]\,,$$
$$\tilde{k}_2(t, t') = \sigma^2[R(t - t') - R(t - t_0)\,R(t_0 - t')]\,. \tag{3.26}$$

Setting $t = t'$ in the last formula, we obtain the conditional variance

$$[\mathrm{D}\xi(t)]_{cond} = \sigma^2[1 - R^2(t - t_0)]\,. \tag{3.27}$$

Note that the a posteriori process is nonstationary. As $t - t_0 \to 0$, the statistical scatter of the values of $\xi(t)$ becomes progressively smaller, since $R(t - t_0) \to 1$ as $t - t_0 \to 0$. As $|t - t_0| \to \infty$, the correlation coefficient $R(t - t_0) \to 0$, and any difference between the a priori and a posteriori processes disappears.

3. Non-Gaussian Processes and Quasi-Moment Functions

When the higher-order correlation functions $k_3(t_1, t_2, t_3)$, $k_4(t_1, t_2, t_3, t_4)$, ... are different from zero, we have a non-Gaussian (or non-normal) random process. In this case, the expression (3.1) is no longer sufficient, and in (1.67) we must also take account of terms for which $s > 2$. Although the characteristic function of a non-Gaussian process can immediately be written as an expansion of the form (1.67), the calculation of the corresponding probability densities is difficult, since in general, the Fourier transform of (1.67) cannot be calculated directly.

In order to simplify the problem of finding the multidimensional distributions of a non-Gaussian process, we introduce the so-called *quasi-moment functions*[3]

$$b_3(t_1, t_2, t_3)\,, \qquad b_4(t_1, t_2, t_3, t_4), \ldots \tag{3.28}$$

[3] P. I. Kuznetsov, R. L. Stratonovich and V. I. Tikhonov, *Quasi-moment functions in the theory of random processes*, Dokl. Akad. Nauk SSSR, **94**, 615 (1954); Theory of Prob. and Its Appl., English edition, **5**, 80 (1960).

These functions occupy an intermediate position (so to speak) between the correlation functions and the moment functions, and, together with $k_1(t_1)$ and $k_2(t_1, t_2)$, they also completely characterize the process. By definition, the quasi-moment functions are connected with the correlation functions by the relation

$$\exp\left\{\sum_{s=3}^{\infty}\frac{i^s}{s!}\sum_{\alpha,\ldots,\omega=1}^{n}k_s(t_\alpha, \ldots, t_\omega)\,z_\alpha \ldots z_\omega\right\} = 1 + \sum_{s=3}^{\infty}\frac{i^s}{s!}\sum_{\alpha,\ldots,\omega=1}^{n}b_s(t_\alpha, \ldots, t_\omega)\,z_\alpha \ldots z_\omega\,, \tag{3.29}$$

which holds for any values of $n, t_1, \ldots, t_n, z_1, \ldots, z_n$. Using (3.29) and (1.67), we can express the characteristic function in terms of quasi-moment functions:

$$\Theta_n(u_1, \ldots, u_n; t_1, \ldots, t_n) = \exp\left\{i\sum_{\alpha=1}^{n}k_1(t_\alpha)\,u_\alpha + \frac{i^2}{2}\sum_{\alpha,\beta=1}^{n}k_2(t_\alpha, t_\beta)\,u_\alpha u_\beta\right\} \times\left\{1 + \sum_{s=3}^{\infty}\frac{i^s}{s!}\sum_{\alpha,\ldots,\omega=1}^{n}b_s(t_\alpha, \ldots, t_\omega)\,u_\alpha \ldots u_\omega\right\}. \tag{3.30}$$

Formula (3.29) allows us to find explicit formulas relating the correlation functions and the quasi-moment functions. In fact, these formulas are the same as (1.69) if we set $k_1(t) \equiv 0$ and $k_2(t_1, t_2) \equiv 0$. The lower-order quasi-moment functions b_3, b_4, b_5 coincide with the corresponding correlation functions. The difference between the correlation functions and the quasi-moment functions begins with b_6:

$$\begin{aligned} b_3 &= k_3\,, \quad b_4 = k_4\,, \quad b_5 = k_5\,, \\ b_6 &= k_6 + 10\{k_3k_3\}_s\,, \\ b_7 &= k_7 + 35\{k_3k_4\}_s\,. \\ &\cdots\quad\cdots\cdots \end{aligned} \tag{3.31}$$

Here, the symbol $\{\ldots\}_s$ denotes the symmetrizing operation, just as in (1.69).

To calculate the n-dimensional probability densities, we have to take the Fourier transform of the characteristic function (3.30):

$$w_n(\xi_1, \ldots, \xi_n; t_1, \ldots, t_n)$$

$$= (2\pi)^{-n} \int \ldots \int \exp\left\{-i\sum_{\alpha=1}^{n} \xi_\alpha u_\alpha\right\}\left\{1 + \sum_{s=3}^{\infty} \frac{i^s}{s!} \sum_{\alpha,\ldots,\omega=1}^{n} b_s(t_\alpha, \ldots, t_\omega)\, u_\alpha \ldots u_\omega\right\} \tag{3.32}$$

$$\times \exp\left\{i\sum_{\alpha=1}^{n} k_1(t_\alpha)\, u_\alpha + \frac{i^2}{2!}\sum_{\alpha,\beta=1}^{n} k_2(t_\alpha, t_\beta)\, u_\alpha u_\beta\right\} du_1 \ldots du_n .$$

Interchanging the operations of summation with respect to s and integration, we can write the last integral in the form

$$w_n(\xi_1, \ldots, \xi_n; t_1, \ldots, t_n) \tag{3.33}$$

$$= \left\{1 + \sum_{s=3}^{\infty} \frac{1}{s!} \sum_{\alpha,\ldots,\omega=1}^{n} b_s(t_\alpha, \ldots, t_\omega)\left(-\frac{\partial}{\partial \xi_\alpha}\right) \ldots \left(-\frac{\partial}{\partial \xi_\omega}\right)\right\} w_n^0(\xi_1, \ldots, \xi_n) ,$$

where w_n^0 is the probability density of a Gaussian process with the same functions $k_1(t)$, $k_2(t, t')$ as the original non-Gaussian process. According to formula (3.3), we have

$$w_n^0(\xi_1, \ldots, \xi_n) = (2\pi)^{-1/2} [\det \| k_2(t_\alpha, t_\beta) \|]^{-1/2} \exp\left\{-\tfrac{1}{2}\sum_{\alpha,\beta=1}^{n} a_{\alpha\beta} x_\alpha x_\beta\right\} , \tag{3.34}$$

where

$$x_\alpha = \xi_\alpha - k_1(t_\alpha) .$$

We can also write the probability density (3.33) in terms of *generalized Hermite polynomials*. Suppose we are given a quadratic form

$$\varphi[x] = \tfrac{1}{2}\sum_{\alpha,\beta=1}^{n} a_{\alpha\beta} x_\alpha x_\beta \tag{3.35}$$

in several variables $x_1, x_2, \ldots, x_n$. Then, with the coefficient matrix $\| a_{\alpha\beta} \|$ we can associate the generalized Hermite polynomials

$$H_\alpha[x] , \qquad H_{\alpha\beta}[x] , \ldots, \qquad H_{\alpha\ldots\omega}[x], \ldots, \tag{3.36}$$

defined by the formula

$$H_{\alpha\ldots\omega}[x] = e^{\varphi[x]}\left(-\frac{\partial}{\partial x_\alpha}\right)\ldots\left(-\frac{\partial}{\partial x_\omega}\right)e^{-\varphi[x]}. \tag{3.37}$$

If we write

$$y_\alpha = y_\alpha[x] = \frac{\partial\varphi[x]}{\partial x_\alpha} = \sum_{\beta=1}^{n} a_{\alpha\beta}a_\beta, \tag{3.38}$$

then, according to (3.37), we have

$$\begin{aligned}
H_\alpha[x] &= y_\alpha,\\
H_{\alpha\beta}[x] &= y_\alpha y_\beta - a_{\alpha\beta},\\
H_{\alpha\beta\gamma}[x] &= y_\alpha y_\beta y_\gamma - a_{\alpha\beta}y_\gamma - a_{\alpha\gamma}y_\beta - a_{\beta\gamma}y_\alpha,\\
H_{\alpha\beta\gamma\delta}[x] &= y_\alpha y_\beta y_\gamma y_\delta - a_{\alpha\beta}y_\gamma y_\delta - a_{\alpha\gamma}y_\beta y_\delta - a_{\alpha\delta}y_\beta y_\gamma - a_{\beta\gamma}y_\alpha y_\delta\\
&\quad - a_{\beta\delta}y_\alpha y_\gamma - a_{\gamma\delta}y_\alpha y_\beta - a_{\alpha\beta}a_{\gamma\delta} - a_{\alpha\gamma}a_{\beta\delta} - a_{\alpha\delta}a_{\beta\gamma},\\
&\quad\ldots\ldots\ldots
\end{aligned} \tag{3.39}$$

In terms of the generalized Hermite polynomials, we can write the probability density (3.33) in the form

$$\begin{aligned}
&w_n(\xi_1, \ldots, \xi_n; t_1, \ldots, t_n)\\
&= \left\{1 + \sum_{s=3}^{\infty}\frac{1}{s!}\sum_{\alpha,\ldots,\omega=1}^{n} b_s(t_\alpha, \ldots, t_\omega)\, H_{\alpha\ldots\omega}[\xi(t) - k_1(t)]\right\} w_n^0(\xi_1, \ldots, \xi_n).
\end{aligned} \tag{3.40}$$

Thus, we see that from a knowledge of the quasi-moment functions, we can immediately write down the probability distributions of any order as an expansion involving generalized Hermite polynomials. The greatest difficulty encountered in doing this is finding the matrix $\| a_{\alpha\beta} \|$ which is the inverse of the correlation matrix $\| k_2(t_\alpha, t_\beta) \|$. Only the first few terms of the expansion (3.40) play an important role, and if we restrict ourselves to a certain number of terms of low order, we can neglect the terms of higher order.

In the special case where $n = 1$, we have a one-dimensional

probability density. When the indices coincide, the polynomials (3.39) become ordinary Hermite polynomials, i.e.,

$$H_{\underbrace{1\ldots1}_{m\text{ times}}}[x] = a_{11}^{m/2} H_m(a_{11}^{-1/2} y_1) = a_{11}^{m/2} H_m(a_{11}^{1/2} x_1)\,, \tag{3.41}$$

where

$$H_m(z) = e^{z^2/2}\left(-\frac{d}{dz}\right)^m e^{-z^2/2}\,.$$

If we bear in mind that

$$a_{11} = \frac{1}{k_2(t,t)} = \frac{1}{\sigma^2}$$

in the one-dimensional case, then, according to (3.40) and (3.41), we find that the probability density is

$$w_1[\xi(t)] = \left\{1 + \sum_{s=3}^{\infty} \frac{1}{s!}\frac{b_s}{\sigma^s} H_s\left[\frac{\xi(t) - k_1(t)}{\sigma}\right]\right\} w^0[\xi(t)]\,, \tag{3.42}$$

where

$$b_s = b_s(t, \ldots, t)\,, \qquad w^0(\xi) = \frac{1}{\sqrt{2\pi}\,\sigma}\exp\left\{-\frac{1}{2\sigma^2}[\xi - k_1(t)]^2\right\}.$$

The expansion (3.42) is known as *Edgeworth's series.*

Next, we consider the two-dimensional case ($n = 2$). Then, when the process is stationary, the matrix $\| a_{\alpha\beta} \|$ is given by (3.7), and the probability density w_2^0 has the form (3.8). In this case, the variables y_1 and y_2 appearing in (3.39) are

$$y_1 = \frac{1}{\sigma^2}\frac{x_1 - Rx_2}{1 - R^2}\,, \qquad y_2 = \frac{1}{\sigma^2}\frac{-Rx_1 + x_2}{1 - R^2}\,. \tag{3.43}$$

In the two-dimensional case, the indices in (3.39) can only take the values 1 and 2. If we write

$$H_{\underbrace{1\ldots1}_{l\text{ times}}\underbrace{2\ldots2}_{m\text{ times}}} = H_{(lm)}\,, \tag{3.44}$$

then it follows from (3.39) that

$$
\begin{aligned}
H_{11} &= H_{(20)} = y_1^2 - a_{11}\,, \\
H_{12} &= H_{(11)} = y_1 y_2 - a_{12}\,, \\
H_{22} &= H_{(02)} = y_2^2 - a_{22}\,, \\
H_{111} &= H_{(30)} = y_1^3 - 3a_{11}y_1\,, \\
H_{112} &= H_{(21)} = y_1^2 y_2 - 2a_{12}y_1 - a_{11}y_2\,, \\
H_{122} &= H_{(12)} = y_1 y_2^2 - 2a_{12}y_2 - a_{22}y_1\,, \\
H_{222} &= H_{(03)} = y_2^3 - 3a_{22}y_2\,.
\end{aligned}
\tag{3.45}
$$

We can write (3.45) as

$$
\begin{aligned}
&H_{(20)} = \frac{1}{\sigma^3\mu^2}(z_1^2 - 1), \quad H_{(11)} = \frac{1}{\sigma^2\mu^2}(z_1 z_2 + R), \quad H_{(02)} = \frac{1}{\sigma^3\mu^2}(z_2^2 - 1), \\
&H_{(30)} = \frac{1}{\sigma^3\mu^3}(z_1^3 - z_1), \qquad H_{(21)} = \frac{1}{\sigma^3\mu^3}(z_1^2 z_2 + 2Rz_1 - z_2), \\
&H_{(12)} = \frac{1}{\sigma^3\mu^3}(z_1 z_2^2 + 2Rz_2 - z_1), \qquad H_{(03)} = \frac{1}{\sigma^3\mu^3}(z_2^3 - z_2),
\end{aligned}
\tag{3.46}
$$

if we introduce the notation

$$
\begin{aligned}
&1 - R^2 = \mu^2\,, \\
&z_1 = \sigma\mu y_1 = \frac{x_1 - Rx_2}{\sigma\mu}\,, \qquad z_2 = \sigma\mu y_2 = \frac{-Rx_1 + x_2}{\sigma\mu}\,.
\end{aligned}
\tag{3.47}
$$

Combining terms in the expansion (3.40) which differ only by permutations of the indices 1 and 2, we can write the two-dimensional probability density as

$$
w_2(\xi_1, \xi_2) = \left\{1 + \sum_{s=3}^{\infty} \frac{1}{s!} \sum_{l+m=s} b_{(lm)} H_{(lm)}[\xi(t) - k_1(t)]\right\} w_2^0(\xi_1, \xi_2)\,,
\tag{3.48}
$$

where

$$b_{(lm)} = b_{l+m}(\overbrace{t_1, \ldots, t_1}^{l \text{ times}}, \overbrace{t_2, \ldots, t_2}^{m \text{ times}}).$$

An expansion of this type is essentially an expansion in orthogonal polynomials.

As we have just seen, the quasi-moment functions are the coefficients in an expansion of the probability density in an infinite series of generalized Hermite polynomials. This is true for any number of random variables $\xi(t_1), \ldots, \xi(t_n)$ corresponding to values of the original process. If the points $t_1, \ldots, t_n$ are made arbitrarily close together, we obtain the corresponding expression for the probability functional.

CHAPTER 4

Markov Processes and Related Processes

1. Definition of a Markov Process. The Stochastic Equation

Markov processes, or processes without aftereffect, are a convenient abstraction. Although the actual processes encountered in radio engineering are not exactly Markovian, it is sometimes a good approximation to regard them as being Markovian. This makes it possible to obtain many concrete results by using the effective mathematical methods of Markov process theory. We begin by considering ideal point processes without aftereffect. Then, we look into the question of when, and in what sense, an actual process encountered in radio engineering can be regarded as a Markov process.

Let $x(t)$ be a random process, and let $x(t_1), \ldots, x(t_n)$ be a set of its values at the consecutive time instants $t_1 > t_2 > \ldots > t_n$. Consider the conditional probability density of the value of $x(t)$ at the most recent time t_1:

$$w[x(t_1) \mid x(t_2), \ldots, x(t_n)] = \frac{w_n[x(t_1), \ldots, x(t_n)]}{w_{n-1}[x(t_2), \ldots, x(t_n)]}. \tag{4.1}$$

The process $x(t)$ is said to be a *Markov process* (or a *process without aftereffect*) if the conditional probability density (4.1) depends only on the last value $x(t_2)$ and not on the preceding values $x(t_3), \ldots, x(t_n)$ $[t_2 > t_3, \ldots, t_2 > t_n]$. Of course, (4.1) can still depend on $x(t_1)$, t_1, t_2, and hence for a Markov process, we can write

$$w(x_1 \mid x_2, \ldots, x_n) = p_{t_1 t_2}(x_1, x_2) \qquad (n > 2), \tag{4.2}$$

where $x_1 = x(t_1), \ldots, x_n = x(t_n)$.

If we bear in mind that by definition, the conditional probabilities (4.1) corresponding to different values of n are always connected by the *compatibility relations*

$$\int w(x_1 \mid x_2, \ldots, x_n)\, w(x_2, \ldots, x_n)\, dx_n = w(x_1 \mid x_2, \ldots, x_{n-1})\, w(x_2, \ldots, x_{n-1}), \tag{4.3}$$

then it can be seen by substituting (4.2) into (4.3) that the same function $p_{t_1 t_2}(x_1, x_2)$ appears in (4.2) for all n. This function equals the conditional probability density

$$p_{t_1 t_2}(x_1, x_2) = w(x_1 \mid x_2) \qquad (t_1 > t_2), \tag{4.4}$$

and is called the *transition probability*. The definition of conditional probability implies the formula

$$w(x_1, \ldots, x_n) = w(x_1 \mid x_2, \ldots, x_n)\, w(x_2, \mid x_3, \ldots, x_n) \ldots w(x_{n-1} \mid x_n)\, w(x_n), \tag{4.5}$$

from which it follows by using (4.2) and (4.4) that in the case of Markov processes, the multidimensional probability densities factor into products of transition probabilities, i.e.,

$$w(x_1, \ldots, x_n) = p_{t_1 t_2}(x_1, x_2)\, p_{t_2 t_3}(x_2, x_3) \ldots p_{t_{n-1} t_n}(x_{n-1}, x_n)\, w(x_n) \ldots \tag{4.6}$$

Thus, if we know the univariate probability distribution and transition probability of $x(t)$, we can write any multivariate probability distribution of $x(t)$, i.e., the functions $p_{tt'}(x, x')$ and $w[x(t)]$ completely characterize a process without aftereffect.

The transition probability has to satisfy certain conditions. In the first place, there is the normalization condition

$$\int p_{tt'}(x, x')\, dx = 1, \tag{4.7}$$

which follows from (4.4), say. Moreover, integrating the distribution (4.6) with respect to some intermediate value $x(t_k)$, $1 < k < n$, we obtain a probability density w_{n-1} of lower order, which in turn satisfies an equation like (4.6). In fact, we have the

following integral equation, called the *Smoluchowski equation* (or the *Chapman-Kolmogorov equation*):

$$\int p_{t_1 t_2}(x_1, x_2)\, p_{t_2 t_3}(x_2, x_3)\, dx_2 = p_{t_1 t_3}(x_1, x_3) \qquad (t_1 > t_2 > t_3). \tag{4.8}$$

To obtain (4.8), we need only consider the relation

$$\int w_3[x(t_1), x(t_2), x(t_3)]\, dx(t_2) = w_2[x(t_1), x(t_3)]\,, \tag{4.9}$$

and then write w_2 and w_3 in the form (4.6).

Next, we study the time dependence of the one-dimensional probability density. Setting $n = 2$ in (4.6) and integrating with respect to x_2, we obtain the equation

$$w(x_1, t_1) = \int p_{t_1 t_2}(x_1, x_2)\, w(x_2, t_2)\, dx_2\,, \tag{4.10}$$

where

$$w(x_1, t_1) = w[x(t_1)]\,, \qquad t_1 > t_2\,.$$

To convert (4.10) into a differential equation, we choose t_1 close to t_2. Setting

$$t_2 = t\,, \qquad t_1 = t + \tau\,, \qquad x_2 = x\,, \qquad x_1 = x_\tau\,,$$

we can write (4.10) in the form

$$w_\tau(x_\tau) = \int p_{t+\tau, t}(x_\tau, x)\, w(x)\, dx\,. \tag{4.11}$$

We now introduce the characteristic function

$$\Theta(u; x) = \langle e^{iu(x_\tau - x)} \rangle = \int e^{iu(x_\tau - x)}\, p_{t+\tau, t}(x_\tau, x)\, dx_\tau \tag{4.12}$$

of the random increment $x_\tau - x$ which occurs during the time interval $[t, t + \tau]$, given that $x(t) = x$. Substituting the inverse transform

$$p_{t+\tau, t}(x_\tau, x) = \frac{1}{2\pi} \int e^{-iu(x_\tau - x)}\, \Theta(u; x)\, du \tag{4.13}$$

into (4.11), we obtain

$$w_\tau(x_\tau) = \frac{1}{2\pi}\int\int e^{-iu(x_\tau - x)}\,\Theta(u;x)\,du\,w(x)\,dx\,. \tag{4.14}$$

According to the usual formula (1.22), the characteristic function (4.12) equals

$$\Theta(u;x) = 1 + \sum_{s=1}^{\infty}\frac{(iu)^s}{s!}\,m_s(x) \tag{4.15}$$

in terms of the moments

$$m_s(x) = \langle (x_\tau - x)^s\rangle \tag{4.16}$$

of the increment $x_\tau - x$. It follows that

$$w_\tau(x_\tau) = \sum_{s=0}^{\infty}\frac{1}{s!}\frac{1}{2\pi}\int\int e^{-iu(x_\tau - x)}\,(iu)^s\,du\,m_s(x)\,w(x)\,dx\,. \tag{4.17}$$

However, since

$$\begin{aligned}\frac{1}{2\pi}\int e^{-iu(x_\tau - x)}\,(iu)^s\,du &= \left(-\frac{\partial}{\partial x_\tau}\right)^s\frac{1}{2\pi}\int e^{-iu(x_\tau - x)}\,du\\ &= \left(-\frac{\partial}{\partial x_\tau}\right)^s\delta(x_\tau - x)\,,\end{aligned} \tag{4.18}$$

we find that

$$w_\tau(x_\tau) = w(x_\tau) + \sum_{s=1}^{\infty}\frac{1}{s!}\left(-\frac{\partial}{\partial x_\tau}\right)^s[m_s(x_\tau)\,w(x_\tau)]\,. \tag{4.19}$$

Dividing by τ and passing to the limit $\tau \to 0$, we obtain

$$\dot{w}(x) = \sum_{s=1}^{\infty}\frac{1}{s!}\left(-\frac{\partial}{\partial x}\right)^s[K_s(x)\,w(x)]\,, \tag{4.20}$$

where the overdot denotes differentiation with respect to t, and

$$K_s(x) = \lim_{\tau\to 0}\frac{m_s(x)}{\tau}\,, \tag{4.21}$$

provided that the limits (4.21) exist.[1] Equation (4.20) is called the *stochastic* (or *kinetic*) *equation*. In the case where the sum of derivatives in the right-hand side of (4.20) has an infinite number of terms, the sum is equivalent to an integral operator, but otherwise (4.20) is a partial differential equation.

According to (4.21), the moments $m_s(x) = \langle (x_\tau - x)^s \rangle$ depend on τ in the following way:

$$\begin{aligned} m_1(x) &= K_1(x)\,\tau + O(\tau^2)\,, \\ m_2(x) &= K_2(x)\,\tau + O(\tau^2)\,, \\ m_3(x) &= K_3(x)\,\tau + O(\tau^2)\,, \\ &\cdots\cdots \end{aligned} \tag{4.22}$$

Using the formulas (1.24), we can easily obtain completely analogous relations for the cumulants $k_s(x)$ of the random increment $x_\tau - x$:

$$k_s(x) = K_s(x)\,\tau + O(\tau^2) \qquad (s = 1, 2, \ldots)\,. \tag{4.23}$$

If we define $\xi(t'; x)$ as the derivative

$$\xi(t'; x) \equiv \xi(t') = \frac{dx(t')}{dt'} \qquad (t' > t) \tag{4.24}$$

of the conditional random process $x(t')$ which takes the value $x(t) = x$, then

$$x_\tau - x = \int_t^{t+\tau} \xi(t')\,dt'\,, \tag{4.25}$$

and the cumulants $k_s(x)$ can be expressed in terms of the correlation functions $k_s(t_1, \ldots, t_s)_\xi$ of the process $\xi(t)$:

$$k_s(x) = \int_t^{t+\tau} \cdots \int_t^{t+\tau} k_s(t_1, \ldots, t_s)_\xi\,dt_1 \ldots dt_s \qquad (s = 1, 2, \ldots)\,. \tag{4.26}$$

---

[1] It should be noted that (4.19) and (4.20) have a simple operator interpretation. For example, (4.19) can be written in the form

$$w_\tau = \Theta\left(i\frac{\partial}{\partial x}, x\right) w\,,$$

if we agree that the second operator x in the function Θ of the operators $i(\partial/\partial x)$ and x always acts before the operator $i(\partial/\partial x)$.

Comparing (4.26) and (4.23), we find that the correlation functions $k_s(t_1, ..., t_s)_\xi$ have delta-functions singularities of the form

$$k_s(t_1, ..., t_s)_\xi = K_s(x)\,\delta(t_1 - t_2) \dots \delta(t_1 - t_s) + A_s \qquad (s = 2, 3, ...), \tag{4.27}$$

where the A_s denote other possible functions which are less singular than the first term (with spectra that fall off faster as the frequency is increased). If we substitute (4.27) into (4.26) and integrate, the functions A_s give rise to terms of order $\tau^2, \tau^3, ...$ The functions $K_s(x)$ are called the *intensity coefficients* of the derivative (4.24), which correspond to the given value $x(t) = x$ at the given time t. These intensity coefficients appear in the stochastic equation (4.20), and in nontrivial cases, at least one of them is different from zero. As the above considerations show, Markov processes are intimately related to certain special random processes whose correlation functions contain delta functions. Such random processes will be said to be *delta-correlated.*

2. Delta-Correlated Processes

Suppose we have a delta-correlated random process $\xi(t)$, described by the correlation functions

$$k_s(t_1, ..., t_s) = K_s(t_1)\,\delta(t_1 - t_2) \dots \delta(t_1 - t_s) \qquad (s = 2, 3, ...), \tag{4.28}$$

where the $K_s(t)$ are the intensity coefficients, which in general depend on the time. In the special case where the process $\xi(t)$ is stationary, the coefficients K_s are constant. Setting $t_1 = t_2 = ... = t_s$ in (4.28) we see at once that a delta-correlated process has infinite cumulants. This shows that delta-correlated processes are not encountered in radio engineering practice. Nevertheless, such processes are sometimes useful as approximate or auxiliary working concepts.

A special role is played by stationary Gaussian delta-correlated processes. Let $\xi(t)$ be such a process, with zero mean value. Then,

among the functions (4.28), only the second-order correlation function

$$k_2(t_1 - t_2) = K_2\delta(t_1 - t_2) \tag{4.29}$$

is nonzero, i.e., $k_1 = K_3 = K_4 = \ldots = 0$, and according to (4.29) and (2.13), $\xi(t)$ has the constant spectral density

$$S[\xi; \omega] = 2K, \qquad \kappa(\omega) = K_2 = K. \tag{4.30}$$

Therefore, $\xi(t)$ will sometimes be called *Gaussian white noise*, or simply *white noise*, by analogy with white light, which has a constant spectral density.

The spectral density of any process encountered in practice is only constant over a certain range, and there is always some cutoff frequency ω_c which is an upper bound for the frequencies appearing in the spectral density of the process. This means that the formula [cf. (2.15)]

$$\mathbf{D}\xi = \frac{1}{\pi}\int_0^\infty \kappa(\omega)\, d\omega \tag{4.31}$$

does not give an infinite value for the variance of the process, but rather a finite value, whose order of magnitude is $K\omega_c/\pi \sim K/\tau_{cor}$. Replacing the actual process by a delta-correlated process means that this frequency ω_c is not explicitly taken into account. This is permissible if ω_c is considerably larger than all other frequencies which are important for the given system or problem. On the other hand, the correlation time $\tau_{cor} = 1/\omega_c$ will then be small, in fact, smaller than all other relevant time constants of the system. Under these conditions, we can make the substitution

$$k_2(t_2 - t_1) \to \delta(t_2 - t_1)\int_{-\infty}^\infty k_2(\tau)\, d\tau, \tag{4.32}$$

where the coefficient of the delta function is chosen in such a way that both sides of (4.32) give the same result when integrated with respect to $t_2 - t_1$.

3. The Fokker-Planck Equation and the Kolmogorov Equation

A Markov process is said to be *continuous* if its higher-order intensity coefficients K_3, K_4, ... equal zero. In this case, equation (4.20) takes the form

$$\dot{w}(x) = -\frac{\partial}{\partial x}[K_1(x)\,w(x)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[K_2(x)\,w(x)]\,, \tag{4.33}$$

and is called the *Fokker-Planck equation* (or the *diffusion equation*).[2] Introducing the *probability current*

$$G(x) = K_1(x)\,w(x) - \frac{1}{2}\frac{\partial}{\partial x}[K_2(x)\,w(x)]\,, \tag{4.34}$$

we can write the Fokker-Planck equation in the form

$$\dot{w} + \frac{\partial G}{\partial x} = 0\,, \tag{4.35}$$

which can be interpreted as the *equation of conservation of probability*. The probability current (4.34) describes the amount of probability crossing the abscissa x in the positive direction per unit time. Consider the interval $x_1 \leqslant x \leqslant x_2$, with end points x_1 and x_2. Then $G(x_1)\tau$ is the amount of probability entering this interval in time τ across the abscissa x_1, and $G(x_2)\tau$ is the amount of probability leaving the interval in time τ across the abscissa x_2. If probability never disappears inside the interval $x_1 \leqslant x \leqslant x_2$, and if there are no sources of probability inside $x_1 \leqslant x \leqslant x_2$, then

[2] Concerning the physical meaning of the Fokker-Planck equation, see e.g., L. S. Pontryagin, A. A. Andronov and A. A. Vitt, *On the statistical analysis of dynamical systems*, Zh. Eksper. Teor. Fiz., 3, 165 (1933), reprinted in A. A. Andronov, *Selected Works*, Izd. Akad. Nauk SSSR, Moscow (1956), p. 142; H. A. Kramers, *Brownian motion in a field of force and the diffusion model of chemical reactions*, Physica, 7, 284 (1940); P. I. Kuznetsov, R. L. Stratonovich and V. I. Tikhonov, *Correlation functions in the theory of Brownian motion. Generalization of the Fokker-Planck equation*, Zh. Eksper. Teor. Fiz., 26, 189 (1954).

the difference $G(x_1)\tau - G(x_2)\tau$ is the increment of the total probability $\int_{x_1}^{x_2} w(x)\,dx$ in the interval $x_1 \leqslant x \leqslant x_2$, i.e.,

$$\int_{x_1}^{x_2} w_\tau(x)\,dx - \int_{x_1}^{x_2} w(x)\,dx = G(x_1)\,\tau - G(x_2)\,\tau\,.$$

Dividing this equation by τ and $x_2 - x_1$, and passing to the limits $\tau \to 0$, $x_2 - x_1 \to 0$, we obtain the equation of conservation (4.35). Thus, just as the equation of diffusion or heat conduction involves a flow of mass or heat, the equation (4.35) involves a flow of probability. In fact, from a mathematical point of view, all these phenomena are described by an equation like (4.35) or (4.33).

In every particular realization of a Markov process, the trajectory $x(t)$ has a very complicated form, and can be thought of as being swept out by a "representative point," which moves along the x-axis just like a Brownian particle or a particle undergoing diffusion. If we take a large number of realizations of the random process, we obtain a large number of "representative points," which move about in a random and erratic fashion. These points form a kind of "gas" which undergoes diffusion, and whose density at any point is proportional to the probability density. Each separate point represents a "molecule" of the "gas," and as they move about, "molecules" cannot be created or destroyed. Thus, in this model, we can talk about the "number of moving points" instead of the "amount of probability," and we can treat $\int_{x_1}^{x_2} w(x)\,dx$ as the number of points in the interval $x_1 \leqslant x \leqslant x_2$, or as the amount of time that a single moving point spends in this interval. Of course, in this language, we have to talk about the *relative* number of points in an interval (rather than the total number), or the *relative* amount of time spent in an interval by a single point, since these are the quantities corresponding to probabilities.

In order to obtain solutions of the Fokker-Planck equation (4.33), we have to supplement it with initial conditions and boundary conditions. If we specify an arbitrary initial distribution

$$w(x, t_0) = w_0(x)$$

at some initial time t_0, we can find the subsequent evolution of this

distribution [i.e., the function $w(x, t)$ for $t > t_0$] by solving equation (4.33). If the initial distribution is a delta function, so that $w_0(x) = \delta(x - x_0)$, then the resulting probability density is just the transition probability $p_{tt_0}(x, x_0)$. Therefore, the transition probability can be found as the solution of the equation

$$\frac{\partial p_{tt_0}(x, x_0)}{\partial t} = -\frac{\partial}{\partial x}[K_1(x)\, p_{tt_0}(x, x_0)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[K_2(x)\, p_{tt_0}(x, x_0)], \quad (4.36)$$

with the initial condition

$$p_{t_0t_0}(x, x_0) = \delta(x - x_0). \quad (4.37)$$

It can be shown[3] that regarded as a function of x_0 and t_0, the transition probability $p_{tt_0}(x, x_0)$ satisfies another differential equation

$$\frac{\partial p_{tt_0}(x, x_0)}{\partial t_0} = -\frac{\partial p_{tt_0}(x, x_0)}{\partial x_0} K_1(x_0) - \frac{1}{2}\frac{\partial^2 p_{tt_0}(x, x_0)}{\partial x_0^2} K_2(x_0). \quad (4.38)$$

This equation, which is the adjoint of equation (4.36), is called the *Kolmogorov equation*. In the case of a stationary random process, $K_1(x)$ and $K_2(x)$ do not depend on t, and the transition probability $p_{t+\tau,t} = p_\tau$ depends only on τ and not on t. Then

$$\frac{\partial p_{tt_0}}{\partial t_0} = -\frac{\partial p_{tt_0}}{\partial t} = -\frac{\partial p_\tau}{\partial \tau},$$

and the Kolmogorov equation (4.38) gives a second expression for the derivative $\partial p_\tau/\partial\tau$, i.e.,

$$\frac{\partial p_\tau}{\partial \tau} = \frac{\partial p_\tau(x, x_0)}{\partial x_0} K_1(x_0) + \frac{1}{2}\frac{\partial^2 p_\tau(x, x_0)}{\partial x_0^2} K_2(x_0), \quad (4.39)$$

in addition to the expression for $\partial p_\tau/\partial\tau$ given by (4.36).

---

[3] See e.g. A. T. Bharucha-Reid, *Elements of the Theory of Markov Processes and Their Applications*, McGraw-Hill Book Co., Inc., New York (1960), Chap. 3. The Kolmogorov equation is often called the *backward equation*, and the Focker-Planck equation the *forward equation*, for an evident reason.

Next, we consider the boundary conditions which have to be taken into account when solving the Fokker-Planck equation (4.33) or (4.36). If the function $x(t)$ can take all possible values from $-\infty$ to ∞, the Fokker-Planck equation is valid on the whole infinite x-axis, and then the boundary conditions take the form of conditions at $\pm\infty$. Integrating (4.35) with respect to x from $-\infty$ to ∞, and bearing in mind that the normalization condition $\int w(x)\,dx = 1$ is satisfied identically for all t, we easily see that the condition

$$G(-\infty, t) = G(\infty, t) \tag{4.40}$$

must hold. However, in addition to (4.40), the stronger conditions

$$G(-\infty, t) = G(\infty, t) = 0 \tag{4.41}$$

and

$$w(-\infty, t) = w(\infty, t) = 0 \tag{4.42}$$

are usually satisfied. According to (4.41), the "representative points" of the process cannot appear at infinity or leave at infinity. In cases where the function $x(t)$ can only take bounded values lying in some interval

$$x_1 \leqslant x \leqslant x_2\,, \tag{4.43}$$

we only consider the Fokker-Planck equation in this region, and the boundary conditions take the form

$$G(x_1, t) = 0\,, \qquad G(x_2, t) = 0\,. \tag{4.44}$$

According to (4.44), there is no flow of "representative points" across the boundary, i.e., no random trajectory can enter the region (4.43) by crossing the boundary, and every random trajectory terminates when it arrives at the boundary.

Of course, depending on the problem under consideration, we can have other boundary conditions as well. For example, in Vol. II, Chap. 9, we shall consider a case where the boundary condition takes the form of a periodicity condition $w(x + 2\pi, t) = w(x, t)$. The question of boundary conditions is also approached differently in first-passage time problems, which will be discussed below.

4. Solutions of the Fokker-Planck Equation

The probability density $w(x, t)$ satisfying the Fokker-Planck equation (4.33) is uniquely determined by the prescribed initial and boundary conditions. If the coefficients $K_1(x)$ and $K_2(x)$ do not depend on the time, then, as time increases, the distribution $w(x, t)$ usually approaches a stationary distribution $w_{st}(x)$, which does not depend on the initial distribution $w(x, t_0)$ or on the time t:

$$\dot{w}_{st}(x) = 0 .$$

Setting $w_{st} = 0$ in (4.35), we find that

$$G_{st}(x) = G_{st} = \text{const} , \tag{4.45}$$

i.e., the probability current is a constant, independent of both x and t. According to (4.34), for a fixed value of G_{st}, equation (4.45) is a linear differential equation in w_{st}:

$$\frac{\partial}{\partial x}[K_2(x)\, w_{st}(x)] - 2K_1(x)\, w_{st}(x) = -2G_{st} . \tag{4.46}$$

The general solution of (4.46) can be obtained by ordinary methods. If we write $K_2(x)\, w_{st}(x) = v(x)$, then (4.46) becomes

$$\frac{\partial v}{\partial x} - 2\frac{K_1}{K_2} v = -2G_{st} . \tag{4.47}$$

Multiplying (4.47) by

$$\exp\left\{-2\int_{x_1}^{x} \frac{K_1(y)}{K_2(y)}\, dy\right\}$$

and integrating with respect to x, we find that (4.47) has the solution

$$v(x) = -2G_{st}\int_{x_1}^{x} \exp\left\{2\int_{x'}^{x} \frac{K_1}{K_2}\, dy\right\} dx' + C\exp\left\{2\int_{x_1}^{x} \frac{K_1}{K_2}\, dy\right\},$$

from which it follows that

$$w_{st}(x) = \frac{C}{K_2(x)} \exp\left\{2\int_{x_1}^{x} \frac{K_1}{K_2}\, dy\right\} - \frac{2G_{st}}{K_2(x)}\int_{x_1}^{x} \exp\left\{2\int_{x'}^{x} \frac{K_1}{K_2}\, dy\right\} dx' . \tag{4.48}$$

Here, C is an arbitrary constant of integration, and the limit of integration x_1 is also arbitrary. However, apart from G_{st}, there is essentially only one arbitrary constant in the solution (4.48), since changing x_1 is equivalent to changing C. The choice of the constant C (or equivalently, of x_1) is determined by the normalization condition, while the probability current G_{st} is found from the boundary conditions. If we choose (4.41) or (4.44) as the boundary conditions, then the expression (4.48) simplifies to

$$w_{st}(x) = \frac{C}{K_2(x)} \exp\left\{2 \int_{x_1}^{x} \frac{K_1(y)}{K_2(y)}\right\} dy . \tag{4.49}$$

Thus, from a knowledge of the functions $K_1(x)$ and $K_2(x)$, we can immediately write down the stationary distribution in terms of a quadrature. This shows the effectiveness of stochastic methods based on the Fokker-Planck equation. Unfortunately, investigation of "transient processes," which involves calculation of the transition probability by solving the nonstationary equation (4.33), is a much more difficult problem. However, in certain special cases, the problem can be simplified, e.g., when $K_2(x)$ is a constant and $K_1(x) = ax + b$ is a linear function of x; then, the process is Gaussian and can be studied by using the theory of Gaussian processes.

In some instances, the nonstationary Fokker-Planck equation can be solved by applying the method of separation of variables, i.e., we look for a solution of the form

$$w(x, t) = X(x)\, T(t) .$$

Then, dividing both sides of equation (4.33) by $w(x, t)$, we obtain

$$\frac{\dot{T}(t)}{T(t)} = \left\{-\frac{\partial}{\partial x}[K_1(x)\, X(x)] + \frac{1}{2}\frac{\partial_2}{\partial x^2}[K_2(x)\, X(x)]\right\} X^{-1}(x) = -\lambda , \tag{4.50}$$

where λ is a constant. For certain prescribed boundary conditions (usually zero boundary conditions), the resulting equation

$$\frac{1}{2}\frac{\partial^2}{\partial x^2}[K_2(x)\, X(x)] - \frac{\partial}{\partial x}[K_1(x)\, X(x)] + \lambda X(x) = 0 \tag{4.51}$$

has solutions consisting of a sequence of eigenfunctions $X_0(x)$, $X_1(x)$, $X_2(x)$, ..., with corresponding eigenvalues λ_0, λ_1, λ_2, ... Integrating (4.51) with respect to x, we easily obtain

$$\lambda_m \int X_m(x)\, dx = 0\,, \tag{4.52}$$

if the difference in the currents

$$G[X_m] = -\frac{1}{2}\frac{\partial}{\partial x}[K_2 X_m] + K_1 X_m$$

evaluated at the limits of integration (i.e., at the boundaries) vanishes. It follows from (4.52) that either $\lambda_m = 0$ or $\int X_m(x)\, dx = 0$. The value $\lambda = 0$ is actually an eigenvalue, and in fact, for $\lambda = 0$, equation (4.51) reduces to the stationary equation. Thus, the eigenfunction X_0 corresponding to $\lambda_0 = 0$ is just the stationary distribution (4.48) or (4.49) found above, i.e.,

$$X_0(x) = w_{st}(x)\,. \tag{4.53}$$

The other eigenfunctions X_m satisfy the condition

$$\int X_m(x)\, dx = 0\,. \tag{4.54}$$

In terms of the eigenfunctions, we can write the solution of equation (4.33) in the form

$$w(x, t) = T_0 X_0(x) + \sum_{m=1}^{\infty} T_m e^{-\lambda_m (t-t_0)} X_m(x)\,, \tag{4.55}$$

where the coefficients T_m are calculated by using the initial conditions. To calculate T_m, it is convenient to use the orthogonality of the eigenfunctions (with weight $1/w_{st}$), as expressed by the formula

$$\int X_m(x)\, X_n(x) \frac{dx}{w_{st}(x)} = \delta_{mn}\,. \tag{4.56}$$

(It is assumed that $\lambda_m \neq \lambda_n$ if $m \neq n$.)

★ To establish the orthogonality relation (4.56), we write (4.51) for two different eigenfunctions:

$$\frac{\partial}{\partial x} G[X_m] = \lambda_m X_m , \qquad \frac{\partial}{\partial x} G[X_n] = \lambda_n X_n$$

where

$$G[X] = -\frac{1}{2}\frac{\partial}{\partial x}[K_2 X] + K_1 X .$$

We multiply the first of these equations by X_n/w_{st} and the second by X_m/w_{st}, and then integrate with respect to x. The result is

$$\int X_n \frac{\partial}{\partial x} G[X_m] \frac{dx}{w_{st}} = \lambda_m \int X_n X_m \frac{dx}{w_{st}} , \tag{4.57}$$

$$\int X_m \frac{\partial}{\partial x} G[X_n] \frac{dx}{w_{st}} = \lambda_n \int X_m X_n \frac{dx}{w_{st}} . \tag{4.58}$$

Next, we prove that the left-hand sides of equations (4.57) and (4.58) are equal. Consider the integral

$$I = \int g \frac{\partial}{\partial x} G[w_{st} f]\, dx = -\frac{1}{2}\int g \frac{\partial^2}{\partial x^2}[K_2 w_{st} f]\, dx + \int g \frac{\partial}{\partial x}[K_1 w_{st} f] ,$$

which can be written in the form

$$I = -\frac{1}{2}\int g K_2 w_{st} \frac{\partial^2 f}{\partial x^2}\, dx - \int g \frac{\partial}{\partial x}[K_2 w_{st}] \frac{\partial f}{\partial x}\, dx - \frac{1}{2}\int g f \frac{\partial^2}{\partial x^2}[K_2 w_{st}]\, dx$$
$$+ \int g K_1 w_{st} \frac{\partial f}{\partial x}\, dx + \int g f \frac{\partial}{\partial x}[K_1 w_{st}]\, dx .$$

The third and fifth terms vanish because of the fact that the stationary distribution w_{st} satisfies the equation

$$\frac{\partial G[w_{st}]}{\partial x} = 0 .$$

Moreover, it is clear from (4.46) that under the condition

$$G[w_{st}] = 0 , \tag{4.59}$$

the second term equals

$$-2\int gK_1 w_{st}\frac{\partial f}{\partial x}\,dx\,,$$

and can be combined with the fourth term. Therefore, we have

$$I = -\frac{1}{2}\int gK_2 w_{st}\frac{\partial^2 f}{\partial x^2}\,dx - \int gK_1 w_{st}\frac{\partial f}{\partial x}\,dx\,.$$

Integrating the last equation by parts, we obtain

$$I = -\frac{1}{2}\int \frac{\partial^2}{\partial x^2}[K_2 w_{st} g]f\,dx + \int \frac{\partial}{\partial x}[K_1 w_{st} g]f\,dx = \int \frac{\partial}{\partial x}G[w_{st} g]f\,dx\,,$$

provided that

$$g\frac{\partial f}{\partial x} - \frac{\partial g}{\partial x}f = 0$$

or

$$gG[w_{st} f] - G[w_{st} g]f = 0 \tag{4.60}$$

on the boundary of the region under consideration.

If we now set $g = X_n/w_{st}$, $f = X_m/w_{st}$, we see that the relation

$$\int X_n \frac{\partial}{\partial x}G[X_m]\frac{dx}{w_{st}} = \int X_m \frac{\partial}{\partial x}G[X_n]\frac{dx}{w_{st}} \tag{4.61}$$

holds, i.e., the left-hand sides of (4.57) and (4.58) are equal, as asserted. For this to be true, it is sufficient that the condition (4.59) and at least one of the boundary conditions

$$G[X] = 0 \quad \text{or} \quad \frac{X}{w_{st}} = 0 \quad \text{on the boundary}$$

be satisfied. It follows from (4.57), (4.58) and (4.61) that

$$(\lambda_m - \lambda_n)\int X_m X_n \frac{dx}{w_{st}} = 0\,,$$

which proves that the eigenfunctions corresponding to different

eigenvalues are orthogonal, with weight $1/w_{st}$. Here, we assume that each eigenfunction is normalized in such a way that

$$\int X_m^2 \frac{dx}{w_{st}} = 1 \qquad (m = 0, 1, 2, \ldots).$$

Thus, we have proved the orthogonality relation (4.56) in the case where there exists a stationary distribution w_{st} satisfying (4.59), and no two eigenfunctions have the same eigenvalue. In most cases encountered in practice, the system of functions $X_m(x)$ is complete.★

The coefficients T_m appearing in the expansion (4.55) can now be expressed in terms of the initial distribution $w(x, t_0)$. Setting $t = t_0$ in (4.55), multiplying both sides of the resulting equation by X_m/w_{st}, and integrating with respect to x, we obtain

$$T_m = \int w(x, t_0)\, X_m(x) \frac{dx}{w_{st}(x)}, \tag{4.62}$$

where we have used the orthogonality relation (4.56). If the initial distribution is a delta function, i.e., if

$$w(x, t_0) = \delta(x - x_0),$$

then the distribution $w(x, t)$ is just the transition probability $p_{tt_0}(x, x_0)$, and in this case, according to (4.55) and (4.62), we have

$$p_{tt_0}(x, x_0) = \sum_{m=0}^{\infty} \frac{X_m(x)\, X_m(x_0)}{w_{st}(x_0)}\, e^{-\lambda_m(t-t_0)}. \tag{4.63}$$

By using the transition probability (4.63) and the initial distribution, we can write probability densities of any order. When the initial distribution is stationary, the expression for the two-dimensional probability density $w_\tau(x, x_0) = p_{t,t-\tau}(x, x_0) w_{st}(x_0)$ is especially concise:

$$w_\tau(x, x_0) = \sum_{m=0}^{\infty} X_m(x)\, X_m(x_0)\, e^{-\lambda_m|\tau|}. \tag{4.64}$$

In particular, this implies that the correlation function and spectral density of the random process $y(t) = F[x(t)]$ are given by

$$k_y(\tau) = \sum_{m=1}^{\infty} h_m^2 e^{-\lambda_m|\tau|} \tag{4.65}$$

and

$$S[y; \omega] = \sum_{m=1}^{\infty} \frac{4\lambda_m h_m^2}{\omega^2 + \lambda_m^2}, \tag{4.66}$$

where

$$h_m = \int F(x)\, X_m(x)\, dx\,. \tag{4.67}$$

Example 1. Consider the equation

$$\dot{w} = \frac{\partial}{\partial x}(\beta x w) + \frac{K}{2}\frac{\partial^2 w}{\partial x^2}\,.$$

In this case, (4.51) takes the form

$$\sigma^2 \frac{\partial^2 X}{\partial x^2} + \frac{\partial}{\partial x}(xX) + \frac{\lambda}{\beta} X = 0 \qquad \left(\sigma^2 = \frac{K}{2\beta}\right). \tag{4.68}$$

If we impose zero boundary conditions at $x = \pm\infty$, then, according to formulas (7.355.3) and (7.351) of Ryshik and Gradstein's handbook,[4] equation (4.68) has eigenvalues $\lambda_n/\beta = n$ $(n = 0, 1, 2, \ldots)$ and eigenfunctions proportional to

$$F^{(n+1)}\left(\frac{x}{\sigma}\right),$$

where

$$F^{(n+1)}(z) = \frac{1}{\sqrt{2\pi}} \frac{d^n}{dz^n} e^{-z^2/2}\,.$$

Thus, if we make the appropriate choice of normalization constants,

$$X_n(x) = \frac{1}{\sqrt{n!}} \frac{1}{\sigma} F^{(n+1)}\left(\frac{x}{\sigma}\right), \qquad \lambda_n = n\beta\,, \tag{4.69}$$

and (4.64) is identical with the expansion (3.14), where $R = e^{-\beta|\tau|}$.

[4] I. M. Ryshik and I. S. Gradstein, *Tables of Series, Products, and Integrals*, VEB Deutscher Verlag der Wissenschaften, Berlin (1957), pp. 405, 404, 246.

Next, we consider equation (4.68) for $x \geqslant 0$ (instead of $-\infty < x < \infty$), with the boundary condition $G[X]_{x=0} = 0$. In this case, we can only keep the eigenfunctions (4.69) which have even indices, i.e.,

$$X_m(x) = \frac{2}{\sqrt{(2m)!}} \frac{1}{\sigma} F^{(2m+1)}\left(\frac{x}{\sigma}\right), \qquad \lambda_m = 2m\beta, \tag{4.70}$$

and the expansion (4.64) now becomes

$$w_\tau(x, x_0) = \frac{2}{\sigma^2} \sum_{m=0}^{\infty} \frac{1}{(2m)!} F^{(2m+1)}\left(\frac{x}{\sigma}\right) F^{(2m+1)}\left(\frac{x_0}{\sigma}\right) e^{-2m\beta|\tau|}. \tag{4.71}$$

Example 2. Consider the equation

$$\dot{w} = \frac{\partial}{\partial x}\left[\left(\beta x - \frac{K}{2x}\right) w\right] + \frac{K}{2} \frac{\partial^2 w}{\partial x^2},$$

which describes a *Rayleigh process* (see Chap. 7, Sec. 3). The corresponding equation for the eigenfunctions is

$$\sigma^2 \frac{\partial^2 X}{\partial x^2} + \frac{\partial}{\partial x}(xX) - \sigma^2 \frac{\partial}{\partial x}\left(\frac{X}{x}\right) + \frac{\lambda}{\beta} X = 0 \qquad \left(\sigma^2 = \frac{K}{2\beta}\right),$$

which becomes

$$\frac{\partial^2 X}{\partial z^2} + \frac{\partial X}{\partial z} + \left[\frac{\frac{1}{2} + (\lambda/2\beta)}{z} + \frac{1}{4z^2}\right] X = 0, \tag{4.72}$$

after we make the change of variables $z = x^2/2\sigma^2$. If we impose zero boundary conditions at $x = 0$ and $x = \infty$, then, according to formulas (7.302) and (7.142.1) of the handbook cited above,[5] equation (4.72) has eigenvalues $\lambda_n = 2n\beta$ $(n = 0, 1, 2, \ldots)$ and eigenfunctions proportional to

$$z^{1/2} e^{-z} L_n(z),$$

where

$$L_n(z) = e^z \frac{d^n}{dz^n}(z^n e^{-z})$$

[5] I. M. Ryshik and I. S. Gradstein, *op. cit.*, pp. 396, 381, 247.

are the Laguerre polynomials. Therefore, we have

$$X_n(x) = \frac{1}{n!}\frac{x}{\sigma^2}e^{-x^2/2\sigma^2}L_n\left(\frac{x^2}{2\sigma^2}\right), \tag{4.73}$$

and the expansion (4.64) takes the form

$$w_\tau(x, x_0) = \frac{xx_0}{\sigma^4}e^{-(x^2+x_0^2)/2\sigma^2}\sum_{n=0}^{\infty}\frac{1}{(n!)^2}L_n\left(\frac{x^2}{2\sigma^2}\right)L_n\left(\frac{x_0^2}{2\sigma^2}\right)e^{-2n\beta|\tau|}. \tag{4.74}$$

Example 3. Consider the somewhat more general equation

$$\dot{w} = \frac{\partial}{\partial x}\left[\left(\beta x - \frac{\gamma}{x}\right)w\right] + \frac{K}{2}\frac{\partial^2 w}{\partial x^2}, \tag{4.75}$$

with zero boundary conditions at $x = 0$ and $x = \infty$. The corresponding equation for X is

$$\sigma^2\frac{\partial^2 X}{\partial x^2} + \frac{\partial}{\partial x}(xX) - \sigma^2\frac{2\gamma}{K}\frac{\partial}{\partial x}\left(\frac{X}{x}\right) + \frac{\lambda}{\beta}X = 0,$$

which takes the form

$$\frac{\partial^2 u}{\partial z^2} + \frac{\partial u}{\partial z} + \left[\frac{\frac{1}{2}+\mu+(\lambda/2\beta)}{z} + \frac{\frac{1}{4}-\mu^2}{z^2}\right]u = 0, \tag{4.76}$$

where $4\mu = (2\gamma/K) - 1$, after we make the change of variables

$$z = \frac{x^2}{2\sigma^2}, \qquad u = z^{-\mu}X.$$

Again using formulas (7.302) and (7.142.1) of Ryshik and Gradstein's tables, we find that equation (4.76) has eigenvalues $\lambda_n = 2n\beta$ and eigenfunctions $u_n(z)$ proportional to

$$z^{\mu+\frac{1}{2}}e^{-z}L_n^{(2\mu)}(z).$$

Returning to the old variables and normalizing the eigenfunctions, we obtain

$$X_n(x) = \frac{1}{\sqrt{n!\,\Gamma(n+2\mu+1)\,\Gamma(2\mu+1)}}\frac{x^{4\mu+1}}{2^{2\mu}\sigma^{4\mu+2}}e^{-x^2/2\sigma^2}L_n^{(2\mu)}\left(\frac{x^2}{2\sigma^2}\right).$$

Then, according to (4.64), we have the expansion

$$w_t(x, x_0) = \frac{2}{\sigma^2}\left(\frac{xx_0}{2\sigma^2}\right)^{4\mu+1} e^{-(x^2+x_0^2)/2\sigma^2} \sum_{n=0}^{\infty} \frac{e^{-2n\beta|\tau|}}{n!\,\Gamma(n+2\mu+1)\,\Gamma(2\mu+1)} \times L_n^{(2\mu)}\left(\frac{x^2}{2\sigma^2}\right) L_n^{(2\mu)}\left(\frac{x_0^2}{2\sigma^2}\right), \tag{4.77}$$

which is a generalization of the expansions (4.71) and (4.74).

5. The Multidimensional Fokker-Planck Equation

The considerations given in Sec. 1 of this chapter can be immediately generalized to the case of a multidimensional Markov process, which consists of several random functions $x_1(t), \ldots, x_m(t)$. Such a process is described by a *transition probability*

$$p_{tt'}[x_1(t), \ldots, x_m(t); x_1(t'), \ldots x_m(t')],$$

in terms of which we can write the multidimensional probability densities. This transition probability is just the conditional probability density

$$p_{tt'}(x_1, \ldots, x_m; x_1', \ldots, x_m') = \frac{w(x_1, \ldots, x_m; x_1', \ldots, x_m')}{w(x_1', \ldots, x_m')} \tag{4.78}$$

of the random variables $x_1(t) = x_1, \ldots, x_m(t) = x_m$ at a general time t, given the fixed values $x_1(t') = x_1', \ldots, x_m(t') = x_m'$ at a previous time $t' < t$. In the case of a continuous Markov process, the one-dimensional probability density $w[x_1(t), \ldots, x_m(t)]$ satisfies the *multidimensional Fokker-Planck equation*

$$\dot{w}(x_1, \ldots, x_m) = -\sum_{\alpha=1}^{m} \frac{\partial}{\partial x_\alpha}[K_\alpha(x_1, \ldots, x_m)\, w(x_1, \ldots, x_m)] + \frac{1}{2}\sum_{\alpha,\beta=1}^{n} \frac{\partial^2}{\partial x_\alpha \partial x_\beta}[K_{\alpha\beta}(x_1, \ldots, x_n)\, w(x_1, \ldots, x_m)], \tag{4.79}$$

where the *intensity coefficients* K_α and $K_{\alpha\beta}$ are defined by the formulas

$$K_\alpha = \lim_{\tau \to 0} \frac{1}{\tau} \langle x_{\alpha\tau} - x_\alpha \rangle ,$$
$$K_{\alpha\beta} = \lim_{\tau \to 0} \frac{1}{\tau} \langle (x_{\alpha\tau} - x_\alpha)(x_{\beta\tau} - x_\beta) \rangle , \tag{4.80}$$

by analogy with (4.21).

In a special case, which might be called the case of "isotropic fluctuations," the matrix $\| K_{\alpha\beta} \|$ has the form

$$K_{\alpha\beta} = K\delta_{\alpha\beta} , \tag{4.81}$$

and then the Fokker-Planck equation (4.79) can be written in the form

$$\dot{w} = -\sum_{\alpha=1}^{m} \frac{\partial G_\alpha}{\partial x_\alpha} , \tag{4.82}$$

where

$$G_\alpha(x) = K_\alpha(x)\, w(x) - \frac{1}{2} \frac{\partial}{\partial x_\alpha} [K(x)\, w(x)] \tag{4.83}$$

are the components of a probability current vector $\mathbf{G} = (G_1, \ldots, G_m)$ in m-dimensional space (x denotes the set of coordinates $x_1, \ldots, x_m$). As can be seen from (4.82), the stationary probability density w_{st} satisfies the equation

$$\sum_{\alpha=1}^{m} \frac{\partial G_\alpha}{\partial x_\alpha} = 0 . \tag{4.84}$$

However, if $m > 1$, the probability current $\mathbf{G}$ does not have to vanish inside the region R under consideration, even if $\mathbf{G}$ satisfies zero boundary conditions

$$G_\alpha(x) = 0 \qquad (\alpha = 1, \ldots, m) \tag{4.85}$$

on the boundary of R, since *rotational* probability flows can occur.

In fact, the current **G** vanishes in the whole region R, i.e.,

$$K_\alpha(x)\, w_{st}(x) - \frac{1}{2}\frac{\partial}{\partial x_\alpha}[K(x)\, w_{st}(x)] = 0 \tag{4.86}$$

only in a special case which we call the *potential* case.

If we substitute

$$K(x)\, w_{st}(x) = e^{-U(x)} \tag{4.87}$$

into (4.86), we find that

$$\frac{\partial U}{\partial x_\alpha} = -2\frac{K_\alpha}{K}. \tag{4.88}$$

This shows that in the potential case, the quantities

$$\frac{K_\alpha(x)}{K(x)} \qquad (\alpha = 1, \ldots, m)$$

are the components of the gradient of a certain function, and hence satisfy the conditions

$$\frac{\partial}{\partial x_\alpha}\left(\frac{K_\beta}{K}\right) = \frac{\partial}{\partial x_\beta}\left(\frac{K_\alpha}{K}\right). \tag{4.89}$$

If the conditions (4.89) are met, the function U is given by the line integral

$$u(x_1, \ldots, x_m) = -\int_{a_1,\ldots,a_m}^{x_1,\ldots,x_m} \frac{2}{K}(K_1\, dx_1 + \ldots + K_m\, dx_m) + C, \tag{4.90}$$

where $a_1, \ldots, a_m$ and the additive constant C are arbitrary. In terms of the function U, the stationary probability density is given by the formula

$$w_{st}(x_1, \ldots, x_m) = \frac{1}{K(x_1, \ldots, x_m)}\, e^{-U(x_1,\ldots,x_m)}, \tag{4.91}$$

where the additive constant in (4.90) is determined from the normalization condition

$$\int \ldots \int w_{st}(x_1, \ldots, x_m)\, dx_1 \ldots dx_m = 1.$$

The solution (4.90) and (4.91) is an obvious generalization of formula (4.49). If the "potential conditions" (4.89) [i.e., the conditions corresponding to the potential case] are satisfied, we can always write down a stationary distribution (4.91) such that the Fokker-Planck equation is satisfied and the current (4.83) satisfies the zerc boundary conditions (4.85). If the boundary conditions are different, or if the potential conditions (4.89) are not met, the problem of finding the stationary distribution becomes much more complicated.

If the coefficient K does not depend on $x_1, \ldots, x_m$, the conditions (4.89) reduce to the conditions

$$\frac{\partial K_\alpha}{\partial x_\beta} = \frac{\partial K_\beta}{\partial x_\alpha}, \tag{4.92}$$

which are to be expected if we regard the K_α as components of a force which can be derived from a potential.

★ Next, we generalize the above considerations to the case where the matrix $\| K_{\alpha\beta} \|$ is arbitrary, rather than a multiple of the unit matrix, as in (4.81). The absence of probability current is now expressed by the conditions

$$G_\alpha = K_\alpha w_{st} - \frac{1}{2} \sum_{\beta=1}^{m} \frac{\partial}{\partial x_\beta} [K_{\alpha\beta} w_{st}] = 0 . \tag{4.93}$$

Setting $w_{st} = e^{-U}$ in (4.93), we find that

$$\sum_{\beta=1}^{m} K_{\alpha\beta} \frac{\partial U}{\partial x_\beta} = \sum_{\beta=1}^{m} \frac{\partial K_{\alpha\beta}}{\partial x_\beta} - 2K_\alpha . \tag{4.94}$$

Solving these equations for the derivatives $\partial U/\partial x_\beta$, we obtain

$$\frac{\partial U}{\partial x_\gamma} = \sum_\alpha \sum_\beta A_{\gamma\alpha} \frac{\partial K_{\alpha\beta}}{\partial x_\beta} - 2 \sum_\alpha A_{\gamma\alpha} K_\alpha , \tag{4.95}$$

where $\| A_{\gamma\alpha} \|$ is the inverse of the matrix $\| K_{\alpha\beta} \|$, i.e.,

$$\sum_\alpha A_{\gamma\alpha} K_{\alpha\beta} = \delta_{\gamma\beta} .$$

It follows from (4.95) that the potential conditions now take the form

$$\frac{\partial}{\partial x_\delta}\sum_\alpha A_{\gamma\alpha}\left(\sum_\beta \frac{\partial K_{\alpha\beta}}{\partial x_\beta} - 2K_\alpha\right) = \frac{\partial}{\partial x_\gamma}\sum_\alpha A_{\delta\alpha}\left(\sum_\beta \frac{\partial K_{\alpha\beta}}{\partial x_\beta} - 2K_\alpha\right) \quad (\gamma, \delta = 1, \ldots, m). \tag{4.96}$$

If these conditions are met, and if there is no flow of probability through the boundary of the region R, then the assumption that the probability current vanishes everywhere in R is justified, and we can find the stationary probability density by calculating the potential function U from its gradient (4.95). To do so, it is necessary that the matrix $\| K_{\alpha\beta} \|$ be nonsingular so that it has an inverse.★

Regardless of whether the potential conditions are met, the solution of the Fokker-Planck equation can also be simplified in the case where $K_1, \ldots, K_m$ are linear functions of the arguments $x_1, \ldots, x_m$, and the functions $K_{\alpha\beta}$ are independent of $x_1, \ldots, x_m$.

6. First-Passage Time Problems

The apparatus of Markov process theory enables us to solve many problems involving processes which eventually terminate; this category includes *first-passage time problems*, which we shall examine only in the one-dimensional case, where the region under consideration is an interval $x_1 \leqslant x \leqslant x_2$. Let $x(t)$ be a Markov process, and suppose we are given the initial distribution

$$w(x, t_0) = w_0(x) \qquad \left(\int_{x_1}^{x_2} w_0(x)\, dx = 1\right) \tag{4.97}$$

of the values $x(t_0)$. We are interested in the time t_{fp} which it takes the process $x(t)$ to first reach the boundaries $x = x_1$ or $x = x_2$. This so-called *first passage* takes place at times which vary from realization to realization, so that the *first-passage time*

$$\tau_{fp} = t_{fp} - t_0 \tag{4.98}$$

is a random variable.

We now show how to calculate the mean first-passage time $\langle \tau_{fp} \rangle$, confining ourselves to the case where the coefficients K_1 and K_2 in the Fokker-Planck equation (4.33) are independent of time. Excluding from consideration any realization of the random process $x(t)$ as soon as it takes the boundary values x_1 or x_2 for the first time, we describe the remaining realizations by a probability density $w(x, t)$ such that

$$\Delta P = w(x, t)\,\Delta x + O[(\Delta x)^2] \tag{4.99}$$

is the probability that at time t, the process $x(t)$ takes a value in the interval $[x, x + \Delta x]$ without ever having reached the boundary during the entire time interval $[t_0, t]$. Then, the integral

$$W(t) = \int_{x_1}^{x_2} w(x, t)\,dx \tag{4.100}$$

gives the probability that $x(t)$ never reaches the boundary during the time interval $[t_0, t]$. Initially, when no realization has yet managed to reach the boundary, the probability density $w(x, t)$ in (4.99) is the same as the original density (4.97), so that

$$W(t^0) = 1\,. \tag{4.101}$$

At subsequent times, the normalization condition is no longer valid, since more and more realizations are excluded from consideration as a result of having reached the boundary. Sooner or later, all the trajectories arrive at the boundary, and hence

$$W(\infty) = 0\,. \tag{4.102}$$

Inside the interval $[x_1, x_2]$, the behavior of the probability density $w(x, t)$ is described by the usual Fokker-Planck equation (4.33) [or by the conservation equation (4.35)], since trajectories cannot terminate inside $[x_1, x_2]$. In fact, trajectories are excluded from consideration only when the boundary is reached, and there is a nonzero probability current at the boundary, corresponding to a flow of "representative points" which are "absorbed" by the boundary. Therefore, the boundary conditions have to be altered in a basic way, as now described. We have already noted that

the derivative $dx(t)/dt$ of a Markov process has infinite variance. This means that the instantaneous velocity of a "representative point" (i.e., a point describing a trajectory of the process) is infinite. However, only a finite displacement occurs in a finite time, which is explained by the fact the velocity of the representative point changes its sign with "infinite frequency," while the point moves in both directions. Thus, if the random function takes the value x at time t, then in the very recent past it must have taken values both larger than x and smaller than x. Therefore, at time $t > t_0$, there are practically no trajectories near the boundary which have not yet touched the boundary. But it is just these trajectories which are described by the probability density $w(x, t)$, and hence $w(x, t)$ vanishes on the boundary:

$$w(x_1, t) = w(x_2, t) = 0 \qquad (t > t_0)\,. \tag{4.103}$$

Zero boundary conditions like these are typical of problems involving first-passage times.

The initial and boundary conditions (4.97) and (4.103) uniquely determine $w(x, t)$ as a solution of the Fokker-Planck equation (4.33). After calculating $w(x, t)$, we can find the probability

$$W(t_0) - W(t) = 1 - W(t) \tag{4.104}$$

that the boundary is first reached during the time interval $[t_0, t]$. Differentiating (4.104), we obtain the probability density of the first-passage time (4.98):

$$w(\tau_{fp}) = -\frac{dW}{dt}(t_0 + \tau_{fp})\,. \tag{4.105}$$

Using (4.105), we find that the mean first-passage time is

$$\langle \tau_{fp} \rangle = -\int_{\tau_{fp}=0}^{\infty} \tau_{fp}\, dW(t_0 + \tau_{fp}) = \int_{t_0}^{\infty} W(t)\, dt\,, \tag{4.106}$$

where we have integrated by parts and used (4.102). If we choose a fixed coordinate value x_0 as the initial condition, so that

$$w_0(x) = \delta(x - x_0)\,,$$

then $w(x, t)$ is just the transition probability $p_{tt_0}(x, x_0)$, which satisfies equation (4.36) with the boundary condition

$$p_{tt_0}(x, x_0) = 0 \quad \text{for} \quad x = x_1, x_2 . \tag{4.107}$$

Moreover, (4.107) holds for $x_0 = x_1, x_2$, since then the boundary is reached from the very outset, and all realizations are excluded from consideration. The transition probability also satisfies the Kolmogorov equation (4.39).

Next, consider the mean first-passage time

$$M(x_0) = \langle \tau_{fp} \rangle = \int_{t_0}^{\infty} W(t, x_0)\, dt , \tag{4.108}$$

regarded as a function of the initial value x_0, where

$$W(t, x_0) = \int_{x_1}^{x_2} p_{tt_0}(x, x_0)\, dx . \tag{4.109}$$

To find an equation satisfied by $M(x_0)$, we integrate the Kolmogorov equation (4.39) with respect to x from x_1 to x_2. Using (4.109), we find that

$$\frac{\partial W}{\partial t} = K_1(x_0) \frac{\partial W}{\partial x_0} + \frac{1}{2} K_2(x_0) \frac{\partial^2 W}{\partial x_0^2} . \tag{4.110}$$

Then, integrating (4.110) with respect to t from t_0 to ∞, and taking into account (4.108), (4.101) and (4.102), we have

$$-1 = K_1(x_0) \frac{dM}{dx_0} + \frac{1}{2} K_2(x_0) \frac{d^2 M}{dx_0^2} . \tag{4.111}$$

If the initial point x_0 lies at the boundary itself, then the boundary is "reached immediately," and the mean first-passage time is zero. Thus, equation (4.111) satisfies the boundary condition

$$M(x_1) = M(x_2) = 0 . \tag{4.112}$$

We are now in a position to write down the solution of (4.111) in a general form. This allows us to find the mean first-passage time (4.106) for an arbitrary initial distribution (4.97), without having

to solve the nonstationary Fokker-Planck equation (4.33). In fact, substituting the expression

$$w(x, t) = \int_{x_1}^{x_2} p_{tt_0}(x, x_0)\, w_0(x_0)\, dx_0$$

into (4.100) and (4.106), we obtain

$$\langle \tau_{fp} \rangle = \int_{x_1}^{x_2} M(x_0)\, w_0(x_0)\, dx_0\,, \tag{4.113}$$

after using (4.108) and (4.109). Equation (4.113) is the generalization of (4.108) to the case of an arbitrary intial distribution $w_0(x)$. For example, suppose that $K_1 = 0$, $K_2 = \text{const}$. Then, the solution of (4.111) is

$$M(x_0) = -\frac{1}{K_2} x_0^2 + C_1 x_0 + C_2\,,$$

which becomes

$$M(x_0) = \frac{1}{K_2} (x_0 - x_1)(x_2 - x_0)\,, \tag{4.114}$$

after using the boundary conditions (4.112). Of course, in a more general case, the result is more complicated, but it can always be expressed in terms of quadratures. Another approach to problems involving first-passage times is given in Vol. II (Chap. 1, Sec. 3, and Chap. 2).

7. Replacement of an Actual Process by a Markov Process. A Special Case

Any actual process encountered in radio engineering, unlike a process which is exactly Markovian, satisfies certain conditions involving its smoothness, differentiability, etc. Nevertheless, in some cases we can treat actual processes as if they were processes without aftereffect. In this section, we examine under what conditions, and with what justification, this can be done.

Suppose that the process $x(t)$ is due to the action of another process $\xi(t)$ on a system whose behavior can be described by the first-order differential equation[6]

$$\dot{x} = \epsilon F[x, \xi(t)], \tag{4.115}$$

called the *fluctuation equation.* Here, ϵ is a small parameter, and F is a known function of the arguments x and ξ. We shall assume that the process $\xi(t)$ describing the external noise perturbations is stationary and has a finite correlation time τ_{cor}.

We begin by considering the special case where the function F does not explicitly depend on x. Then, since $F[\xi(t)]$ can be regarded as a new random function, there is no loss of generality in writing

$$\dot{x} = \epsilon\xi(t). \tag{4.116}$$

Choosing $t = 0$ as the initial time, we can write the solution of (4.116) in the form

$$x(t) = x(0) + \epsilon \int_0^t \xi(t')\, dt', \tag{4.117}$$

If the initial value $x(0) = x_0$ is not random, the cumulants k_s of $x(t)$ can be obtained by integrating the correlation functions $k_s(t_1, \ldots, t_s)_\xi$ of the process $\xi(t)$:

$$k_s = \epsilon^s \int_0^t \cdots \int_0^t k_s(t_1, \ldots, t_s)_\xi \, dt_1 \ldots dt_s \qquad (s > 1). \tag{4.118}$$

Using the cumulants k_s, we can write the characteristic function

$$\Theta(u, t) = \exp\left\{\sum_{s=1}^{\infty} \frac{(iu)^s}{s!} k_s\right\} \tag{4.119}$$

of the random increment $x(t) - x_0$. Then, taking the time derivative

[6] More generally, F can also involve the time t *explicitly*, in which case equation (4.115) becomes

$$\dot{x} = \epsilon F[x(t), \xi(t), t].$$

of (4.119), we find that the characteristic function satisfies the differential equation

$$\dot{\Theta} = \left[\sum_{s=1}^{\infty} \frac{(iu)^s}{s!} \dot{k}_s\right] \Theta . \tag{4.120}$$

Next, we find the corresponding differential equation for the probability density

$$w(x - x_0, t) = \frac{1}{2\pi} \int e^{-iu(x-x_0)}\, \Theta(u, t)\, du . \tag{4.121}$$

Taking the time derivative of both sides of (4.121), and substituting (4.120) into the integrand in the right-hand side of the resulting equation, we obtain

$$\dot{w} = \sum_{s=1}^{\infty} \frac{\dot{k}_s}{s!} \frac{1}{2\pi} \int (iu)^s\, e^{-iu(x-x_0)}\, \Theta(u, t)\, du . \tag{4.122}$$

Since

$$\frac{1}{2\pi} \int (iu)^s\, e^{-iu(x-x_0)}\, \Theta\, du = \left(-\frac{\partial}{\partial x}\right)^s \frac{1}{2\pi} \int e^{-iu(x-x_0)}\, \Theta\, du$$

$$= \left(-\frac{\partial}{\partial x}\right)^s w ,$$

it follows from (4.122) that

$$\dot{w} = \sum_{s=1}^{\infty} \frac{(-1)^s}{s!} \dot{k}_s \frac{\partial^s w}{\partial x^s} . \tag{4.123}$$

Thus, we have obtained a stochastic equation of the type (4.20) for a process which is not Markovian. Even in the case of a stationary process $\xi(t)$, the coefficients $\dot{k}_s$ depend on the time. However, this dependence is appreciable only for time intervals of the order of the correlation time ($t \sim \tau_{cor}$), and the coefficients $\dot{k}_s$ are practically constant for $t \gg \tau_{cor}$. In fact, according to (4.118) and (2.2), we have

$$\dot{k}_s = \epsilon^s s \int_0^t \cdots \int_0^t k_s(t, t_1, \ldots, t_{s-1})_\xi\, dt_1 \ldots dt_{s-1}$$

$$= \epsilon^s s \int_{-t}^0 \cdots \int_{-t}^0 k'_s(\tau_1, \ldots, \tau_{s-1})_\xi\, d\tau_1 \ldots d\tau_{s-1} .$$

If we assume that the integrals

$$\int_{-\infty}^{0} \dots \int_{-\infty}^{0} k_s'(\tau_1, \dots, \tau_{s-1})\, d\tau_1 \dots d\tau_{s-1} \qquad (s > 1) \tag{4.124}$$

are absolutely convergent, it follows that

$$k_s \to \epsilon^s K_s \quad \text{as} \quad t \to \infty, \tag{4.125}$$

where

$$K_s = s \int_{-\infty}^{0} \dots \int_{-\infty}^{0} k_s'(\tau_1, \dots, \tau_{s-1})\, d\tau_1 \dots d\tau_{s-1} \qquad (s > 1) \tag{4.126}$$

are the *intensity coefficients* of the process $\xi(t)$. It can be shown that the K_s are also given by the expressions

$$\begin{aligned} K_s &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} k_s'(\tau_1, \dots, \tau_{s-1})\, d\tau_1 \dots d\tau_{s-1} \\ &= s! \int_0^{\infty} d\tau_1 \int_0^{\tau_1} d\tau_2 \dots \int_0^{\tau_{s-2}} k_s'(\tau_1, \dots, \tau_{s-1})\, d\tau_{s-1}, \end{aligned} \tag{4.127}$$

where the coefficient K_2 coincides with (2.8):

$$K_2 = K = 2 \int_0^{\infty} k(\tau)\, d\tau.$$

According to (4.125), for sufficiently large time intervals, equation (4.123) takes the form

$$\dot{w} = \sum_{s=1}^{\infty} \frac{(-1)^s}{s!} \epsilon^s K_s \frac{\partial^s w}{\partial x^s}. \tag{4.128}$$

If the parameter ϵ is small, only the lower-order derivatives play an important role, and the higher-order derivatives are unimportant. If we omit terms corresponding to $s > 2$, equation (4.128) turns into the Fokker-Planck equation.

Of course, the fact that the probability density w obeys a stochastic equation does not mean that $x(t)$ is a Markov process. For this to be the case, one would have to be able to express the

multidimensional distributions as products of the transition probabilities, after writing the transition probability itself in terms of the probability density (4.121), i.e.,

$$p_\tau(x, x_0) = w(x - x_0, \tau) .$$

If a multidimensional distribution factors into a product of such functions, this means that the random increments

$$X_1 = x(t_1) - x(0), \qquad X_2 = x(t_2) - x(t_1), \ldots \qquad (0 < t_1 < t_2 < \ldots) , \tag{4.129}$$

corresponding to successive time intervals, are independent of each other. Therefore, the Markov condition is equivalent to the condition that the increments (4.129) be independent. In an actual case, however, where the correlation time is finite, there is correlation between the increments (4.129).

For example, consider the increments

$$\begin{aligned} X_1 &= x(\tau_1) - x(0) \qquad = \epsilon \int_0^{\tau_1} \xi(t)\, dt , \\ X_2 &= x(\tau_1 + \tau_2) - x(\tau_1) = \epsilon \int_{\tau_1}^{\tau_1+\tau_2} \xi(t)\, dt , \\ &\ldots\ldots \qquad\qquad \ldots , \end{aligned} \tag{4.130}$$

corresponding to the adjacent intervals $[0, \tau_1]$ and $[\tau_1, \tau_1 + \tau_2]$. The covariance (or cross correlation) of X_1 and X_2 is given by

$$\mathbf{K}[X_1, X_2] = \epsilon^2 \int_0^{\tau_1} \int_{\tau_1}^{\tau_1+\tau_2} k(t_2 - t_1)\, dt_1\, dt_2 = \epsilon^2 \int_{-\tau_1}^{0} \int_0^{\tau_2} k(t_2 - t_1)\, dt_1\, dt_2 . \tag{4.131}$$

If the integral

$$\int_{-\infty}^{0} \int_0^{\infty} k(t_2 - t_1)\, dt_1\, dt_2 = \int_0^{\infty} \tau k(\tau)\, d\tau \tag{4.132}$$

is absolutely convergent, then, according to (4.131), we have

$$\mathbf{K}[X_1, X_2] = \epsilon^2 \int_0^{\infty} \tau k(\tau)\, d\tau + D , \tag{4.133}$$

where $D \to 0$ as $\tau_1, \tau_2 \to \infty$. It is appropriate to define the correlation time by the formula

$$\tau_{cor} = \frac{\int_0^\infty \tau k(\tau)\, d\tau}{\int_0^\infty k(\tau)\, d\tau} = \frac{2}{K}\int_0^\infty \tau k(\tau)\, d\tau\,, \tag{4.134}$$

which is somewhat different from (2.7), but of the same order of magnitude. Then,

$$\mathbf{K}[X_1, X_2] \to \frac{\epsilon^2}{2} K\tau_{cor} \tag{4.135}$$

as $\tau_1, \tau_2 \to \infty$.

On the other hand, the variance of the random increments X_1 and X_2 is much greater than (4.135). In fact,

$$\begin{aligned}\mathbf{D}X_1 &= \epsilon^2 \int_0^{\tau_1}\int_0^{\tau_1} k(t_1 - t_2)\, dt_1\, dt_2 = 2\epsilon^2 \int_0^{\tau_1} (\tau_1 - \tau)\, k(\tau)\, d\tau \\ &= \epsilon^2 K(\tau_1 - \tau_{cor}) + D_1\,, \qquad (4.136)\\ \mathbf{D}X_2 &= 2\epsilon^2 \int_0^{\tau_2} (\tau_2 - \tau)\, k(\tau)\, d\tau = \epsilon^2 K(\tau_2 - \tau_{cor}) + D_2\,,\end{aligned}$$

where $D_1 \to 0$ as $\tau_1 \to \infty$, and $D_2 \to 0$ as $\tau_2 \to \infty$. It follows from (4.135) and (4.136) that for large τ_1 and τ_2, the correlation coefficient of the increments X_1 and X_2 is approximately

$$R = \frac{\mathbf{K}[X_1, X_2]}{\sigma(X_1)\,\sigma(X_2)} = \frac{1}{2}\frac{\tau_{cor}}{\sqrt{(\tau_1 - \tau_{cor})(\tau_2 - \tau_{cor})}}\,, \tag{4.137}$$

which is small when

$$\tau_1 \gg \tau_{cor}\,, \qquad \tau_2 \gg \tau_{cor}\,. \tag{4.138}$$

Therefore, we can neglect cross correlations, if we only consider intervals with lengths much greater than the correlation time. The same is true of the higher-order correlations, the only difference being that

$$\tau_{cor} = \frac{s!}{K_s}\int_0^\infty \tau_1\, d\tau_1 \int_0^{\tau_1} d\tau_2 \ldots \int_0^{\tau_{s-2}} k'_s(\tau_1, \ldots, \tau_{s-1})\, d\tau_{s-1} \qquad (s > 2) \tag{4.139}$$

should now be chosen as the correlation time.

Thus, for time intervals which are considerably greater than the correlation time, the increments (4.130) can be regarded as completely independent, and the process $x(t)$ can be regarded as a Markov process. This means that the process $\xi(t)$ with correlation functions $k_s(t_1, \ldots t_s)$ can be replaced by a delta-correlated process whose correlation functions are

$$K_s\delta(t_2 - t_1) \ldots \delta(t_s - t_1)\,, \tag{4.140}$$

with the same intensity coefficients

$$K_s = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} k'_s(\tau_1, \ldots, \tau_{s-1})\, d\tau_1 \ldots d\tau_{s-1} \tag{4.141}$$

as the actual process $\xi(t)$.

8. Replacement of an Actual Process by a Markov Process. The General Case

We now consider the more general case where no special restrictions are imposed on the function appearing in the right-hand side of equation (4.115):

$$\dot{x} = \epsilon F[x, \xi(t)] \equiv \epsilon F(x, t)\,. \tag{4.142}$$

Then the derivation of the stochastic differential equation for the one-dimensional probability density is more complicated. Specifying the value

$$x(t_0) = x_0 \tag{4.143}$$

at the initial time t_0, we consider the increment

$$x(t) - x_0 = H(x_0) \equiv H(x_0, t, t_0)\,, \tag{4.144}$$

which now depends in an essential way on the initial value x_0. To find $H(x_0)$, we solve equation (4.142) with the initial condition (4.143). In fact, we look for $H(x_0)$ in the form of an expansion

$$H(x_0) = \epsilon H_1(x_0) + \epsilon^2 H_2(x_0) + \ldots \tag{4.145}$$

Substituting (4.145) into (4.142), we find that

$$\epsilon\dot{H}_1 + \epsilon^2\dot{H}_2 + \ldots = \epsilon F(x_0) + \epsilon \frac{\partial F}{\partial x}(x_0)\,[\epsilon H_1 + \epsilon^2 H_2 + \ldots] + \frac{\epsilon}{2}\frac{\partial^2 F}{\partial x^2}(x_0)\,[\epsilon H_1 + \epsilon^2 H_2 + \ldots]^2 + \ldots$$

Equating terms of the same order in ϵ, we obtain

$$\begin{aligned} \dot{H}_1(x_0) &= F(x_0)\,, \\ \dot{H}_2(x_0) &= \frac{\partial F}{\partial x}(x_0)\,H_1(x_0)\,, \\ \dot{H}_3(x_0) &= \frac{\partial F}{\partial x}(x_0)\,H_2(x_0) + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}(x_0)\,H_1^2(x_0)\,, \\ &\cdots\cdots\cdots \end{aligned} \tag{4.146}$$

so that

$$\begin{aligned} H_1(x_0) &= \int_{t_0}^{t} F(x_0, t')\,dt'\,, \\ H_2(x_0) &= \int_{t_0}^{t} dt' \frac{\partial F}{\partial x}(x_0, t') \int_{t_0}^{t'} F(x_0, t'')\,dt''\,, \end{aligned} \tag{4.147}$$

After we have expressed $H(x_0)$ in terms of F, to some desired accuracy, we consider the statistical characteristics of $H(x_0)$. The characteristic function of $H(x_0)$ can be written in the form

$$\langle e^{iu(x-x_0)}\rangle = 1 + \sum_{s=1}^{\infty} \frac{(iu)^s}{s!}\langle H^s(x_0)\rangle\,. \tag{4.148}$$

By evaluating the inversion integral, we find the one-dimensional probability density

$$w(x \mid x_0) = \frac{1}{2\pi}\int e^{-iu(x-x_0)}\,[1 + \sum_{s=1}^{\infty} \frac{(iu)^s}{s!}\langle H^s(x_0)\rangle]\,du\,. \tag{4.149}$$

Interchanging the order of integration and summation in (4.149), we have

$$w(x \mid x_0) = \left[1 + \sum_{s=1}^{\infty} \frac{1}{s!} \left(-\frac{\partial}{\partial x}\right)^s \langle H^s(x_0)\rangle\right] \frac{1}{2\pi} \int e^{-iu(x-x_0)}\, du$$

$$= \left[1 + \sum_{s=1}^{\infty} \frac{1}{s!} \left(-\frac{\partial}{\partial x}\right)^s (H^s(x_0)\rangle\right] \delta(x - x_0)\,. \tag{4.150}$$

Unlike (4.148), this last expression is of a somewhat formal character, since the sum of derivatives of the delta function does not converge in the ordinary sense. To arrive at expressions that converge in the ordinary sense, we have to integrate the product of (4.150) and a function $f(x_0)$ from some suitable class of functions, obtaining

$$\int w(x \mid x_0) f(x_0)\, dx_0 = f(x) + \sum_{s=1}^{\infty} \frac{1}{s!} \left(-\frac{\partial}{\partial x}\right)^s [\langle H^s(x)\rangle f(x)]\,.$$

We note that here the function $f(x)$ is first multiplied by $\langle H^s(x)\rangle$, and then the product is differentiated.

It is now convenient to introduce the operator

$$L = \sum_{s=1}^{\infty} \frac{1}{s!} \left(-\frac{\partial}{\partial x}\right)^s \langle H^s(x)\rangle\,, \tag{4.151}$$

where in each term, the operation of multiplication precedes that of differentiation. Using L, we can write (4.150) as

$$w = (1 + L)\, \delta(x - x_0)\,. \tag{4.152}$$

Differentiating (4.52) with respect to time, we find that

$$\dot{w} = \dot{L}\delta(x - x_0)\,. \tag{4.153}$$

Using (4.152), we can write the last equation in the form

$$\dot{w} = \dot{L}(1 + L)^{-1} w\,, \tag{4.154}$$

where the operator

$$(1+L)^{-1} = 1 - L + L^2 - L^3 + \ldots \tag{4.155}$$

is the inverse of the operator $1 + L$.

Equation (4.154) is the stochastic equation we are looking for, and the operator $\dot{L}(1+L)^{-1}$ appearing in (4.154) is given in terms of the expansions (4.145), (4.151) and (4.155), which are all power series in the small parameter ϵ, since H and hence L are of order ϵ. Substituting these expansions into $\dot{L}(1+L)^{-1}$ and grouping terms involving the same power of ϵ, we obtain an expansion for $\dot{L}(1+L)^{-1}$, whose first few terms will be explicitly calculated below. As the time interval $t - t_0$ is increased, the convergence of the series (4.145), (4.151) and (4.155) becomes worse. However, the convergence of the series for $\dot{L}(1+L)^{-1}$ does not become worse; instead, this series converges to a limiting series which is independent of the initial time t_0.

We now find the terms of the series for $\dot{L}(1+L)^{-1}$ which are of order ϵ and ϵ^2. Then, in all intermediate calculations, we need only retain those terms which are required for this purpose. According to (4.145), (4.151) and (4.155), we have

$$\begin{aligned}\dot{L} &= \left(-\frac{\partial}{\partial x}\right)\langle \dot{H}\rangle + \left(-\frac{\partial}{\partial x}\right)^2 \langle \dot{H}H\rangle + O(\epsilon^3) \\ &= \left(-\frac{\partial}{\partial x}\right)\langle \epsilon\dot{H}_1 + \epsilon^2\dot{H}_2\rangle + \left(-\frac{\partial}{\partial x}\right)^2 \langle \epsilon^2\dot{H}_1 H_1\rangle + O(\epsilon^3)\end{aligned}$$

and

$$\begin{aligned}(1+L)^{-1} &= 1 - L + O(\epsilon^2) = 1 - \left(-\frac{\partial}{\partial x}\right)\langle H\rangle + O(\epsilon^2) \\ &= 1 - \epsilon\left(-\frac{\partial}{\partial x}\right)\langle H_1\rangle + O(\epsilon^2).\end{aligned}$$

Multiplying $\dot{L}$ and $(1+L)^{-1}$, we obtain

$$\begin{aligned}\dot{L}(1+L)^{-1} = &-\epsilon\frac{\partial}{\partial x}\langle \dot{H}_1 + \epsilon\dot{H}_2\rangle + \epsilon^2\frac{\partial^2}{\partial x^2}\langle \dot{H}_1 H_1\rangle \\ &-\epsilon^2\frac{\partial}{\partial x}\langle \dot{H}_1\rangle\frac{\partial}{\partial x}\langle H_1\rangle + O(\epsilon^3).\end{aligned} \tag{4.156}$$

In the last term of (5.156), we interchange the order of the operators $\partial/\partial x$ and $\langle \dot{H}_1(x) \rangle$. In fact, using the identity

$$\langle \dot{H}_1(x) \rangle \frac{\partial f(x)}{\partial x} = \frac{\partial}{\partial x} [\langle \dot{H}_1(x) \rangle f(x)] - \frac{\partial \langle \dot{H}_1(x) \rangle}{\partial x} f(x),$$

we easily see that

$$\langle \dot{H}_1(x) \rangle \frac{\partial}{\partial x} = \frac{\partial}{\partial x} \langle \dot{H}_1(x) \rangle - \frac{\partial \langle \dot{H}_1(x) \rangle}{\partial x}, \tag{4.157}$$

and hence

$$\frac{\partial}{\partial x} \langle \dot{H}_1 \rangle \frac{\partial}{\partial x} \langle H_1 \rangle = \frac{\partial^2}{\partial x^2} \langle \dot{H}_1 \rangle \langle H_1 \rangle - \frac{\partial}{\partial x} \left\langle \frac{\partial \dot{H}_1}{\partial x} \right\rangle \langle H_1 \rangle .$$

Finally, taking (4.146) into account, we can write the operator (4.156) as

$$\dot{L}(1 + L)^{-1} = -\epsilon \frac{\partial}{\partial x} \langle F \rangle - \epsilon^2 \frac{\partial}{\partial x} \mathbf{K} \left[\frac{dF}{\partial x}, H_1 \right] + \epsilon^2 \frac{\partial^2}{\partial x^2} \mathbf{K}[F, H_1] + O(\epsilon^3). \tag{4.158}$$

By neglecting terms of order ϵ^3 and higher in (4.158), we find that equation (4.154) reduces to the Fokker-Planck equation (4.33), where

$$\begin{aligned} K_1(x) &= \epsilon \langle F(x) \rangle + \epsilon^2 \mathbf{K} \left[\frac{\partial F(x)}{\partial x}, H_1(x) \right], \\ K_2(x) &= 2\epsilon^2 \mathbf{K}[F(x), H_1(x)], \end{aligned} \tag{4.159}$$

or, because of (4.147),

$$K_1(x) = \epsilon \langle F(x) \rangle + \epsilon^2 \int_{t_0 - t}^{0} \mathbf{K} \left[\frac{\partial F(x)}{\partial x}, F_\tau(x) \right] d\tau,$$

$$K_2(x) = 2\epsilon^2 \int_{t_0}^{t} \mathbf{K}[F(x, t), F(x, t')] \, dt' = 2\epsilon^2 \int_{t_0 - t}^{0} \mathbf{K}[F(x), F_\tau(x)] \, d\tau ,$$

where

$$F_\tau(x) = F(x, t + \tau) .$$

As $t - t_0 \to \infty$, we see that

$$K_2(x) \to K(x), \qquad K_1(x) \to M(x) + \tfrac{1}{4} K'(x), \tag{4.160}$$

where

$$M(x) = \epsilon \langle F(x) \rangle$$

$$K(x) = 2\epsilon^2 \int_{-\infty}^{0} \mathbf{K}[F(x), F_\tau(x)]\, d\tau, \tag{4.161}$$

$$K'(x) = 4\epsilon^2 \int_{-\infty}^{0} \mathbf{K}\left[\frac{\partial F(x)}{\partial x}, F_\tau(x)\right].$$

For time intervals $t - t_0$, which are considerably greater than the correlation time τ_{cor} of the function $F[x, \xi(t)]$, or of the function $\xi(t)$, equation (4.154) is practically the same as the Fokker-Planck equation

$$w = -\frac{\partial}{\partial x}\left\{\left[M(x) + \frac{1}{4} K'(x)\right] w\right\} + \frac{1}{2} \frac{\partial^2}{\partial x^2} [K(x)\, w(x)]. \tag{4.162}$$

By analogy with (4.134), it is appropriate to define the correlation time by the formula

$$\tau_{cor} = \frac{2}{K} \epsilon^2 \int_{-\infty}^{0} |\tau|\, \mathbf{K}[F, F_\tau]\, d\tau. \tag{4.163}$$

For time intervals $t_1 - t_0$, $t_2 - t_1$, ... which greatly exceed the correlation time (4.163), the multidimensional probability distributions factor into products of transition probabilities (4.149) satisfying equation (4.162). The reason for this is essentially the same as in the preceding sections, but the proof is more complicated. Then, the random process $x(t)$ will be approximately Markovian, and, to this extent, all that has been said about Markov processes will apply to $x(t)$.

According to (4.21) and (4.16), the coefficient

$$K_1(x) = M(x) + \tfrac{1}{4} K'(x) \tag{4.164}$$

is the "average derivative"

$$\lim_{\tau \to 0} \frac{\langle x_\tau - x \rangle}{\tau} = \langle \dot{x} \rangle = \langle \epsilon F[x(t), \xi(t)] \rangle = K_1(x).$$

It follows that if we wish to calculate the mean value of $\epsilon F[x(t), \xi(t)]$, we have to allow for correlation between $x(t)$ and $\xi(t)$. According to (4.161), as a result of this correlation, $\langle \epsilon F[x(t), \xi(t)] \rangle$ differs from the mean value

$$M(x) = \epsilon \langle F[x, \xi(t)] \rangle , \tag{4.165}$$

calculated without regard for correlation between $x(t)$ and $\xi(t)$, by the quantity[7]

$$\epsilon^2 \int_{-\infty}^{0} \mathbf{K} \left[\frac{\partial F}{\partial x}, F_\tau \right] d\tau + O(\epsilon^3) . \tag{4.166}$$

Similarly, to calculate the mean value of any other function $G[x(t), \xi(t)]$ of the random process $x(t)$ satisfying equation (4.142), subject to the condition $x(t_1) = x_1$, we have to allow for correlation between $x(t)$ and $\xi(t)$ by using the formula

$$\langle G[x(t_1), \xi(t_1)] \rangle = \langle G[x_1, \xi(t_1)] \rangle + \epsilon \int_{-\infty}^{0} \mathbf{K} \left[\frac{\partial G(x_1)}{\partial x}, F_\tau(x_1) \right] d\tau + O(\epsilon^3), \tag{4.167}$$

where the first term in the right-hand side is the mean value calculated without regard for correlation between $x(t)$ and $\xi(t)$.

Among the possible functions $F(x, \xi)$, an important position is occupied by those which satisfy the relation

$$\int_{-\infty}^{0} \mathbf{K} \left[\frac{\partial F}{\partial x}, F_\tau \right] d\tau = \int_{-\infty}^{0} \mathbf{K} \left[F, \frac{\partial F_\tau}{\partial x} \right] d\tau , \tag{4.168}$$

which can be regarded as a kind of time-symmetry condition (i.e., symmetry with respect to the time reflection $t \to -t$). If (4.168) holds, then $K'(x)$ can be obtained by differentiating $K(x)$:

$$K'(x) = \frac{\partial K)x)}{\partial x} . \tag{4.169}$$

[7] When calculating the averages in (4.165) and (4.166), the argument x is regarded as a fixed quantity, rather than as a function of time.

Then equation (4.162) can be written in the form

$$\dot{w} = -\frac{\partial}{\partial x}[Mw] + \frac{1}{4}\frac{\partial}{\partial x}\left[K\frac{\partial w}{\partial x} + \frac{\partial}{\partial x}(Kw)\right]. \qquad (4.170)$$

If M and K are independent of time, (4.170) gives the stationary distribution

$$w_{st}(x) = \frac{C}{\sqrt{K(x)}} \exp\left\{2\int_{x_1}^{x} \frac{M(y)}{K(y)}\,dy\right\}, \qquad (4.171)$$

which is equivalent to (4.49), because of (4.160) and (4.169), and is valid under the same assumptions.

In particular, the relations (4.168) and (4.169) hold in the special case where

$$\dot{x} = \epsilon F[x, \xi(t)] = f(x) + g(x)\,\xi(t)\,; \qquad (4.172)$$

here, $f(x)$ and $g(x)$ are known functions, while $\xi(t)$ is a stationary random perturbation with zero mean value and intensity coefficient κ, i.e.,

$$\langle \xi(t) \rangle = 0\,, \qquad 2\int_{-\infty}^{0} \langle \xi\xi_\tau \rangle\,d\tau = \kappa\,. \qquad (4.173)$$

In this case, according to (4.161)

$$M(x) = f(x)\,, \qquad K(x) = \kappa g^2(x)\,, \qquad (4.174)$$

and the stationary distribution (4.171) becomes

$$w_{st}(x) = \frac{C}{g(x)} \exp\left\{\frac{2}{\kappa}\int_{x_1}^{x} \frac{f(y)}{g^2(y)}\,dy\right\}. \qquad (4.175)$$

We can also solve the inverse problem, which consists in finding a fluctuation equation (4.142) for a random process satisfying a given Fokker-Planck equation. Of course, in general this problem does not have a unique solution. However, the solution will be unique if we restrict ourselves to equations of the type (4.172), containing a *Gaussian* delta-correlated random process $\xi_0(t)$ with zero mean and unit intensity:

$$\langle \xi_0(t) \rangle = 0\,, \qquad \langle \xi_0\xi_{0\tau} \rangle = \delta(\tau)\,, \qquad (\kappa = 1)\,. \qquad (4.176)$$

Then, using (4.164), (4.169) and (4.174), we find that

$$g(x) = \sqrt{K_2(x)}, \qquad f(x) = K_1(x) - \frac{1}{4}\frac{\partial K_2(x)}{\partial x}. \tag{4.177}$$

Therefore, the arbitrary Fokker-Planck equation (4.33) corresponds to the fluctuation equation

$$\dot{x} = K_1(x) - \frac{1}{4}\frac{\partial K_2(x)}{\partial x} + \sqrt{K_2(x)}\,\xi_0(t). \tag{4.178}$$

We can associate a fluctuation equation of the form (4.178) with a more complicated fluctuation equation (4.142) which has the same Fokker-Planck equation; two such equations are said to be (*stochastically*) *equivalent*. This method of simplifying the equation of a random process by replacing it by an equivalent equation turns out to be useful in solving certain problems, and will be used later.

Finally, we consider briefly the error committed when we discard terms of order ϵ^3 and higher in (4.154) and (4.158). Retaining all the terms would lead us to a stochastic equation (4.20), where each coefficient $K_s(x)$ can be expressed in terms of F as a power series in the parameter ϵ (beginning with terms of order ϵ^s), whose terms can be found by applying the method used above. We are mainly interested in the limiting values of the coefficients $K_s(x)$ which come into play when the influence of the initial time t_0 vanishes, i.e., when $t - t_0 \gg \tau_{cor}$. Thus, we go one step further in the direction of making formula (4.159) more precise, by retaining terms of order ϵ^3 in our calculations and neglecting terms of higher orders. This leads[8] to the following formulas:

$$\begin{aligned} K_1 = M + \frac{1}{4}K' + \frac{\epsilon^2}{2}\int_{-\infty}^{0}\int_{-\infty}^{0} \mathbf{K}\left[\frac{\partial^2 F}{\partial x^2}, F_\tau, F_\sigma\right] d\tau\, d\sigma \\ + \epsilon^3 \int_{-\infty}^{0} d\tau \int_{-\infty}^{\tau} \left\{ \mathbf{K}\left[\frac{\partial F}{\partial x}, \frac{\partial F_\tau}{\partial x}, F_\sigma\right] + \left\langle \frac{\partial F_\tau}{\partial x}\right\rangle \mathbf{K}\left[\frac{\partial F}{\partial x}, F_\sigma\right] \right. \\ \left. - \langle F_\tau\rangle \mathbf{K}\left[\frac{\partial F}{\partial x}, \frac{\partial F_\sigma}{\partial x}\right]\right\} d\sigma, \end{aligned}$$

[8] The details are given in the Supplement, p. 126.

$$K_2 = K + 2\epsilon^3 \int_{-\infty}^{0} \int_{-\infty}^{0} \mathbf{K}\left[\frac{\partial F}{\partial x}, F_\tau, F_\sigma\right] d\tau\, d\sigma$$

$$+ 2\epsilon^3 \int_{-\infty}^{0} d\tau \int_{-\infty}^{\tau} \left\{\mathbf{K}\left[F, \frac{\partial F_\tau}{\partial x}, F_\sigma\right] + \left\langle \frac{\partial F_\tau}{\partial x} \right\rangle \mathbf{K}[F, F_\sigma]\right.$$

$$\left. - \langle F_\tau \rangle \mathbf{K}\left[F, \frac{\partial F_\sigma}{\partial x}\right]\right\} d\sigma\,,$$

$$K_3 = 3\epsilon^3 \int_{-\infty}^{0} \int_{-\infty}^{0} \mathbf{K}[F, F_\tau, F_\sigma]\, d\tau\, d\sigma\,,$$

$$K_4 = \ldots = 0\,. \tag{4.179}$$

Here M, K' and K denote the expressions (4.161) previously found, of orders ϵ and ϵ^2. In the special case where the fluctuation equation has the form (4.172), the formulas (4.179) can be considerably simplified and reduce to

$$K_1 = f + \frac{\kappa_2}{2} g \frac{\partial g}{\partial x} + \frac{\kappa_3}{6} g \frac{\partial}{\partial x}\left[g \frac{\partial g}{\partial x}\right] + cg^2 \frac{\partial g}{\partial x} \frac{\partial}{\partial x}\left(\frac{f}{g}\right),$$

$$K_2 = \kappa_2 g^2 + \kappa_2 g^2 \frac{\partial g}{\partial x} + 2cg^3 \frac{\partial}{\partial x}\left(\frac{f}{g}\right), \tag{4.180}$$

$$K_3 = \kappa_2 g^3\,,$$

where

$$\kappa_2 = 2 \int_{-\infty}^{0} \langle \xi \xi_\tau \rangle\, d\tau\,,$$

$$\kappa_3 = 3 \int_{-\infty}^{0} \int_{-\infty}^{0} \mathbf{K}[\xi, \xi_\tau, \xi_\sigma]\, d\tau\, d\sigma\,,$$

$$c = \int_{-\infty}^{0} |\tau| \langle \xi \xi_\tau \rangle\, d\tau\,.$$

We now examine the relative size of the terms appearing in (4.179). Taking account of (4.161), we can make the following order-of-magnitude estimates:

$$\frac{1}{4} K' \sim \epsilon^2 \frac{\partial F}{\partial x} F \tau_{cor}\,,$$

$$K_1 - M - \frac{1}{4} K' \sim \epsilon^3 \frac{\partial^2 F}{\partial x^2} F^2 \tau_{cor}^2 + \epsilon^3 \left(\frac{\partial F}{\partial x}\right)^2 F \tau_{cor}^2. \tag{4.181}$$

Provided that

$$\left(\frac{\partial^2 F}{\partial x^2}\right) F \lesssim \left(\frac{\partial F}{\partial x}\right)^2,$$

we can conclude from (4.181) that the terms in ϵ^3 in the expansion of the coefficient K_1 are smaller by a factor $\epsilon(\partial F/\partial x)\tau_{cor}$ than the term in ϵ^2. Similarly, we can convince ourselves that this factor describes the relative size of the correction $K_2 - K$, compared with K_2. In fact, more generally, the corrections due to higher-order approximations to the coefficients K_s, involving higher powers of the parameter ϵ, have relative sizes described by an appropriate power of the ratio

$$\frac{\tau_{cor}}{\tau_0} \sim \epsilon \frac{\partial F}{\partial x} \tau_{cor}. \tag{4.182}$$

Here,

$$\tau_0 \sim \left[\epsilon \frac{\partial F}{\partial x}\right]^{-1} \tag{4.183}$$

is a time constant, which can be called the *relaxation time.* A necessary condition for the effectiveness of the methods developed in this section is that the inequality

$$\epsilon \frac{\partial F}{\partial x} \tau_{cor} \ll 1 \tag{4.184}$$

should hold.

We introduced the small parameter ϵ in order to make it clearer how to construct successive approximations in determining the coefficients of the stochastic equation. However, in all actual cases, ϵ occurs in combination with the function F. In fact, in solving practical problems, one knows the whole right-hand side $\epsilon F = \Phi$ of equation (4.115), and there is no need to introduce ϵ at all. Thus, the quantity $(\partial\Phi/\partial x)\,\tau_{cor}$ is actually the parameter which determines the convergence of the successive approximations.

9. Systems of Fluctuation Equations Involving Several Processes

All the considerations of the preceding section can be immediately generalized to the case of a system of equations

$$\begin{aligned} x_1 &= \epsilon F_1(x_1, \ldots, x_p, t) = \epsilon F_1(x, t), \\ &\cdots \quad \cdots \quad \cdots \\ x_p &= \epsilon F_p(x_1, \ldots, x_p, t) = \epsilon F_p(x, t), \end{aligned} \tag{4.185}$$

which define the random processes $x_1(t), \ldots, x_p(t)$. For fixed values of $x_1, \ldots, x_p$, the expressions $F_1(x, t), \ldots, F_p(x, t)$ are known random functions of time. Choosing the initial values

$$x_l(t_0) = x_{l0} \qquad (l = 1, \ldots, p), \tag{4.186}$$

we write the solution x_l $(1 \leqslant l \leqslant p)$ of the system (4.185) as an expansion

$$\begin{aligned} x_l(t) - x_{l0} = H_l(x_0) + \epsilon \int_{t_0}^{t} F_l(x_0, t_1)\, dt_1 \\ + \sum_{m=1}^{p} \epsilon^2 \int_{t_0}^{t} dt_1 \frac{\partial F_l}{\partial x_m}(x_0, t_1) \int_{t_0}^{t} F_m(x_0, t_2)\, dt_2 + \ldots \end{aligned} \tag{4.187}$$

The increments $x_1 - x_{10}, \ldots, x_p - x_{p0}$ have the joint characteristic function

$$\begin{aligned} \left\langle \exp\left\{ i \sum_l u_l (x_l - x_{l0}) \right\} \right\rangle = 1 + \sum_l i u_l \langle H_l(x_0) \rangle \\ + \frac{1}{2} \sum_{l,m} i u_l i u_m \langle H_l(x_0)\, H_m(x_0) \rangle + \ldots \end{aligned} \tag{4.188}$$

By analogy with (4.150), we can write the multidimensional probability density, which is the inverse Fourier transform of (4.188), in the form

$$w(x_1 - x_{10}, \ldots, x_p - x_{p0}) = (1 + L)\, \delta(x_1 - x_{10}) \ldots \delta(x_p - x_{p0}), \tag{4.189}$$

where

$$L = \sum_l \left(-\frac{\partial}{\partial x_l} \right) \langle H_l(x) \rangle + \frac{1}{2} \sum_{l,m} \left(-\frac{\partial}{\partial x_l} \right) \left(-\frac{\partial}{\partial x_m} \right) \langle H_l(x)\, H_m(x) \rangle + \ldots \tag{4.190}$$

It follows at once from (4.189) that the probability density $w(x_1 - x_{10}, \ldots, x_p - x_{p0})$ satisfies the equation

$$\dot{w} = \dot{L}(1 + L)^{-1}w\,. \tag{4.191}$$

We must now calculate the operator $\dot{L}(1 + L)^{-1}$. Substituting (4.187) into (4.190), we obtain

$$\begin{aligned} L = &-\epsilon \sum_l \frac{\partial}{\partial x_l} \left\{ \int_{t_0}^{t} \langle F_l(x, t_1)\rangle \, dt_1 \right. \\ &\left. + \epsilon \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} \sum_m \left\langle \frac{\partial F_l(x, t_1)}{\partial x_m} F_m(x, t_2) \right\rangle dt_2 \right\} \\ &+ \frac{\epsilon^2}{2} \sum_{l,m} \frac{\partial^2}{\partial x_l \partial x_m} \int_{t_0}^{t} \int_{t_0}^{t} \langle F_l(x, t_1)\, F_m(x, t_2)\rangle \, dt_1 \, dt_2 + O(\epsilon^2)\,, \end{aligned} \tag{4.192}$$

so that

$$\begin{aligned} \dot{L} = &-\epsilon \sum_l \frac{\partial}{\partial x_l} \left\{ \langle F_l(x, t)\rangle + \epsilon \int_{t_0}^{t} \sum_m \left\langle \frac{\partial F_l}{\partial x_m}(x, t)\, F_m(x, t') \right\rangle dt' \right. \\ &\left. + \epsilon^2 \sum_{l,m} \frac{\partial^2}{\partial x_l \partial x_m} \int_{t_0}^{t} \langle F_l(x, t)\, F_m(x, t')\rangle \, dt' + O(\epsilon^3)\,. \right. \end{aligned} \tag{4.193}$$

Just as before, the operator $\dot{L}(1 + L)^{-1}$ differs from (4.193) only by having the correlation functions

$$\mathbf{K}\left[\frac{\partial F_l}{\partial x_m}, F_m\right], \qquad \mathbf{K}[F_l, F_m]$$

in place of the averages

$$\left\langle \frac{\partial F_l}{\partial x_m} F_m \right\rangle, \qquad \langle F_l F_m\rangle\,.$$

As a result, equation (4.191) becomes

$$\begin{aligned} \dot{w} = &-\epsilon \sum_l \frac{\partial}{\partial x_l} \left\{ \left(\langle F_l\rangle + \epsilon \sum_m \int_{t_0 - t}^{0} \mathbf{K}\left[\frac{\partial F_l}{\partial x_m}, F_{m\tau}\right] d\tau \right) w \right\} \\ &+ \epsilon^2 \sum_{l,m} \frac{\partial^2}{\partial x_l \partial x_m} \left\{ \int_{t_0 - t}^{0} \mathbf{K}[F_l, F_{m\tau}]\, d\tau \, w \right\}, \end{aligned} \tag{4.194}$$

except for terms of order ϵ^3 and higher. Here, in calculating the averages $\langle F_l \rangle$ and the correlation functions in the integrands, the arguments $x_1, \ldots, x_p$ are regarded as fixed. For $t - t_0 \gg \tau_{cor}$, the lower limit $t_0 - t$ can be replaced by $-\infty$.

We see that to within the given accuracy, the equations (4.185) are equivalent to the multidimensional Fokker-Planck equation (4.79). We now show that, conversely, every Fokker-Planck equation can be replaced by a system of fluctuation equations

$$\dot{x}_l = f_l(x) + \sum_{m=1}^{p} g_{lm}(x)\, \xi_m(t) \qquad (l = 1, \ldots, p), \tag{4.195}$$

where the functions $f_l(x)$ and $g_{lm}(x)$ are suitably chosen. Here, $\xi_1(t), \ldots, \xi_p(t)$ are independent, *Gaussian*, delta-correlated random functions with zero mean values and unit intensities:

$$\langle \xi_l \rangle = 0, \qquad \langle \xi_l \xi_{m\tau} \rangle = \delta_{lm}\delta(\tau). \tag{4.196}$$

Applying formula (4.194) to the equations (4.195), using (4.196), and comparing the result with (4.79), we obtain

$$\begin{aligned} K_l(x) &= f_l(x) + \frac{1}{2}\sum_{m,j} \frac{\partial g_{lj}}{\partial x_m} g_{mj}, \\ K_{lm}(x) &= \sum_j g_{lj}(x)\, g_{mj}(x). \end{aligned} \tag{4.197}$$

We can always go from the Fokker-Planck equation to the system (4.195), provided that the relations (4.197) can be solved for the functions $g_{lm}(x)$ and $f_l(x)$. But this can always be done, as we now show. The matrix $K = \| K_{lm} \|$ is symmetric and positive definite (more precisely, nonnegative definite), and hence it follows from familiar results of linear algebra that there exists a real, symmetric, nonnegative definite matrix $G = \| g_{lm} \|$ which is the square root of K, i.e.,

$$G = K^{1/2}, \qquad G^2 = K, \qquad \sum_j g_{lj} g_{jm} = K_{lm}. \tag{4.198}$$

In fact, because of the symmetry condition $g_{jm} = g_{mj}$, (4.198) is the same as the second of the equations (4.197). It follows that the functions g_m, and hence the functions

$$f_i(x) = K_i(x) - \frac{1}{2}\sum_{m,j} \frac{\partial g_{ij}}{\partial x_m} g_{jm} \tag{4.199}$$

can always be determined.

To calculate $\| g_{lm} \|$, we can use the orthogonal matrix $U = \| u_{ij} \|$ which reduces K to diagonal form, i.e.,

$$UKU^{-1} = \| K_i^0 \delta_{ij} \| \qquad (K_i^0 \geqslant 0) . \tag{4.200}$$

In terms of U, we have

$$G = U^{-1} \| \sqrt{K_i^0}\, \delta_{ij} \| \, U . \tag{4.201}$$

Using the u_{ij}, we can write the equations (4.195) in the form

$$\sum_i u_{ii} \dot{x}_i = \sum_i u_{ii} f_i + \sqrt{K_i^0}\, \eta_i(t) , \tag{4.202}$$

where they have been solved for the "fluctuations." Here,

$$\eta_i = \sum_j u_{ij} \xi_j \qquad (i = 1, \ldots, p)$$

are a system of random functions just like the $\xi_i(t)$, and in fact satisfy the same conditions (4.196).

In the *diffusion approximation* (which corresponds to neglecting terms of orders ϵ^3, ϵ^4, ... in the stochastic equation), any system of equations (4.185) can be "simplified," in the sense of being replaced by an equivalent simpler system (4.195) or (4.202).

10. Gaussian Markov Processes

Within the class of Markov processes, the simplest are those which are described by a Fokker-Planck equation (4.79) where the

coefficients K_{lm} are constants and the coefficients $K_l(x)$ are linear functions:

$$K_l(x) = -\sum_m \beta_{lm} x_m + b_l \qquad (l = 1, ..., p). \tag{4.203}$$

According to (4.195) and (4.197), this Fokker-Planck equation is equivalent to the fluctuation equations

$$\dot{x}_l + \sum_m \beta_{lm} x_m = \eta_l(t), \tag{4.204}$$

where

$$\eta_l(t) = b_l + \sum_m g_{lm} \xi_m(t) \tag{4.205}$$

are Gaussian random processes such that

$$\begin{aligned} \langle \eta_l \rangle &= b_l \\ \mathbf{K}[\eta_l, \eta_{m\tau}] &= K_{lm}\delta(\tau). \end{aligned} \tag{4.206}$$

The system of linear equations (4.204) can always be solved, and as a result, $x_1(t)$, ..., $x_p(t)$ can be expressed linearly in terms of the known random processes $\eta_1(t)$, ..., $\eta_p(t)$. Therefore, the processes $x_1(t)$, ..., $x_p(t)$ are also Gaussian, and their statistical characteristics can be easily determined.

In order to calculate the cross-spectral densities of the processes x_1, ..., x_p, we replace the differentiation operator in (4.204) by $i\omega$ and find the inverse of the matrix

$$\| \beta_{lm} + i\omega\delta_{lm} \| .$$

Let $\Delta(i\omega)$ be the determinant of this matrix, and let $\Delta_{lm}(i\omega)$ be the cofactor of the element in row l and column m, so that

$$\| \beta_{lm} + i\omega\delta_{lm} \|^{-1} = \left\| \frac{\Delta_{lm}(i\omega)}{\Delta(i\omega)} \right\|, \tag{4.207}$$

and

$$x_j = \sum_l \frac{\Delta_{jl}(i\omega)}{\Delta(i\omega)} \eta_l . \tag{4.208}$$

Using (4.208), we find that the cross-spectral density of the processes x_j and x_k is

$$S[x_j, x_k; \omega] = \sum_{l,m} \frac{\Delta_{jl}(i\omega)}{\Delta(i\omega)} \frac{\Delta^*_{km}(i\omega)}{\Delta^*(i\omega)} S[\eta_l, \eta_m; \omega]. \qquad (4.209)$$

But because of (4.206), (2.54) and (2.12), we have

$$S[\eta_l, \eta_m; \omega] = 2K_{lm} + 4\pi b_l b_m \delta(\omega).$$

In particular, setting $j = k$, we find that the spectral density of the process $x_j(t)$ [centered at its mean] is given by the expression

$$S[x_j - \langle x_j \rangle; \omega] = 2 \mid \Delta(i\omega) \mid^{-2} \sum_{l,m} \Delta_{jl}(i\omega)\, \Delta^*_{jm}(i\omega)\, K_{lm}. \qquad (4.210)$$

Here, the numerator and denominator contain polynomials in ω^2. A process with a spectral density of this kind is said to have a *rational spectral density.*

Thus, we have shown that Gaussian Markov processes have rational spectral densities. The converse assertion is also true, i.e., every Gaussian process with a rational spectral density can be represented as a component of a multidimensional Markov process. This fact is important from a theoretical point of view, because the spectral density of any actual process can always be approximated to any desired accuracy by a rational function. Therefore, an actual Gaussian process which is non-Markovian can be represented approximately as the component of a multidimensional Markov process, and by increasing the number of components of the Markov process, the approximation can be made more accurate. This question will be discussed further at the end of Sec. 12.

We now consider in more detail the simplest process $x(t)$ with a rational spectral density. This process, which is simultaneously stationary, Gaussian and Markovian, plays an important role in noise theory, and will be called the *exponentially correlated process*, a designation whose meaning will be apparent from the formula derived below for the correlation function of $x(t)$. According to (4.204), for a one-dimensional process we have

$$+ \beta x = \eta, \qquad (4.211)$$

where

$$\langle \eta \rangle = b\,, \qquad \mathbf{K}[\eta, \eta_\tau] = K\delta(\tau)\,.$$

In this case, formulas (4.208) and (4.210) become

$$x = \frac{\eta}{\beta + i\omega}\,,$$

$$S[x - \langle x \rangle; \omega] = \frac{2K}{|\beta + i\omega|^2} = \frac{2K}{\beta^2 + \omega^2}\,. \tag{4.212}$$

To study what happens when the process $x(t)$ is "switched on," we impose the initial condition

$$x(t_0) = x_0\,. \tag{4.213}$$

Then equation (4.211) has the solution

$$x(t) = x_0 e^{-\beta(t-t_0)} + \int_{t_0}^{t} e^{-\beta(t-s)}\,\eta(s)\,ds\,. \tag{4.214}$$

Using (4.214), we find the following expressions for the mean value and the correlation function of $x(t)$:

$$\begin{aligned} \langle x(t) \rangle &= x_0 e^{-\beta(t-t_0)} + \int_{t_0}^{t} e^{-\beta(t-s)}\,b\,ds = \frac{b}{\beta} + \left(x_0 - \frac{b}{\beta}\right) e^{-\beta(t-t_0)}\,, \\ \mathbf{K}[x(t_1), x(t_2)] &= \int_{t_0}^{t_1}\int_{t_0}^{t_2} e^{-\beta(t_1+t_2-s_1-s_2)}\,K\delta(s_1 - s_2)\,ds_1\,ds_2\,. \end{aligned} \tag{4.215}$$

If to be explicit, we choose $t_2 - t_1 \geqslant 0$, then

$$\mathbf{K}[x(t_1), x(t_2)] = K\int_{t_0}^{t_1} e^{-\beta(t_1+t_2-2s)}\,ds = \frac{K}{2\beta}\,[e^{-\beta(t_2-t_1)} - e^{-\beta(t_1+t_2-2t_0)}]\,.$$

In general, we have

$$\mathbf{K}[x(t_1), x(t_2)] = \frac{K}{2\beta}\,[e^{-\beta|t_2-t_1|} - e^{-\beta(t_1+t_2-2t_0)}]\,. \tag{4.216}$$

As $t - t_0$ increases, the "stationary" correlation function

$$k(\tau) = \sigma^2 R(\tau)\,, \qquad \sigma^2 = \frac{K}{2\beta}\,, \qquad R(\tau) = e^{-\beta|\tau|}\,, \tag{4.217}$$

corresponding to the spectral density (4.212), becomes "established." Moreover, according to (4.216), after the process is switched on and while it is reaching equilibrium, its variance is given by the formula

$$\mathbf{D}x(t) = \sigma^2[1 - R^2(t - t_0)]. \tag{4.218}$$

Since the process $x(t)$ is Gaussian, just like $\eta(t)$, the mean value (4.215) and variance (4.218) completely specify the one-dimensional probability density

$$w(x) = p_{t-t_0}(x, x_0) = \frac{1}{\sigma\sqrt{2\pi(1 - R^2)}} \exp\left\{-\frac{[x - m - (x_0 - m)R]^2}{2\sigma^2(1 - R^2)}\right\}, \tag{4.219}$$

where

$$R = R(t - t_0), \qquad m = \frac{b}{\beta}.$$

The probability density $w(x)$ corresponds to the condition (4.213), and hence is just the Markov transition probability (4.4), in terms of which the arbitrary multidimensional distribution (4.6) can be written if we know the initial distribution.

The formulas (4.215) and (4.216) are special cases of the general formulas (3.26). Therefore, the distribution (4.219) gives the conditional density $w[x(t) \mid x_0]$ for any stationary Gaussian process. However, (4.219) can be used to construct the multidimensional distributions only in the case of the exponentially correlated process, for which $R(\tau) = e^{-\beta|\tau|}$.

11. Second-Order Fluctuation Equations. Solutions of Special Cases of the Two-Dimensional Fokker-Planck Equation

In many problems, the noise process $x(t)$ under consideration satisfies a second-order fluctuation equation

$$\ddot{x} = F[x, \dot{x}, \xi(t)], \tag{4.220}$$

where $\xi(t)$ is a random process with a small correlation time τ_{cor}. Introducing the notation

$$\bar{t} = \frac{t}{\tau_{cor}}, \qquad y = \dot{x} = \frac{1}{\tau_{cor}} \frac{dx}{d\bar{t}},$$

we can write (4.220) as

$$\frac{d^2x}{d\bar{t}^2} = \tau_{cor}^2 F\left(x, \frac{1}{\tau_{cor}} \frac{dx}{d\bar{t}}, \xi\right),$$

or as a system

$$\begin{aligned} \frac{dx}{d\bar{t}} &= \tau_{cor}\, y, \\ \frac{dy}{d\bar{t}} &= \tau_{cor} F(x, y, \xi). \end{aligned} \tag{4.221}$$

The small correlation time τ_{cor} in the right-hand side of (4.221) plays the role of the small parameter ϵ which appears in the right-hand side of (4.185) and allows us to go from the system of fluctuation equations (4.185) to the Fokker-Planck equation (4.194). In terms of the original symbols x and $t = \tau_{cor}\bar{t}$, the Fokker-Planck equation takes the form

$$\begin{aligned} \dot{w}(x, \dot{x}) = &- \dot{x} \frac{\partial}{\partial x} w - \frac{\partial}{\partial \dot{x}} \left\{ \left(\langle F \rangle + \int_{-\infty}^{0} \mathbf{K} \left[\frac{\partial F}{\partial \dot{x}}, F_\tau \right] d\tau \right) w \right\} \\ &+ \frac{\partial^2}{\partial \dot{x}^2} \left\{ \int_{-\infty}^{0} \mathbf{K}[F, F_\tau]\, d\tau\, w \right\}. \end{aligned} \tag{4.222}$$

We now analyze some special cases of (4.220).

11.1. Let the fluctuation equation have the form

$$\mu^2 \ddot{x} + \dot{x} - f(x) = \xi(t), \tag{4.223}$$

to which we can reduce the somewhat more general equation

$$a\ddot{x} + b\dot{x} = f(x) + g(\xi). \tag{4.224}$$

Moreover, let $\xi(t)$ be a delta-correlated random process with zero mean, i.e.,

$$\langle \xi \rangle = 0\,, \qquad \langle \xi \xi_\tau \rangle = \kappa \delta(\tau)\,.$$

Then if we write (4.223) as a system

$$\mu \dot{x} = y\,,$$

$$\mu \dot{y} = -\frac{y}{\mu} + f(x) + \xi(t)\,,$$

the corresponding Fokker-Planck equation becomes

$$\dot{w}(x, y) = -\frac{1}{\mu}\left[y\frac{\partial w}{\partial x} + f(x)\frac{\partial w}{\partial y}\right] + \frac{1}{\mu^2}\left[\frac{\partial}{\partial y}(yw) + \frac{\kappa}{2}\frac{\partial^2 w}{\partial y^2}\right], \quad (4.225)$$

from which we can derive the following results:

1. It is easily verified that the stationary distribution corresponding to (4.225) is[9]

$$w_{st}(x, y) = C \exp\left\{-\frac{y^2}{\kappa} + \frac{2}{\kappa}\int_{x_1}^{x} f(z)\,dz\right\}, \quad (4.226)$$

since when w has the form (4.226), each of the terms in brackets in (4.225) vanishes. It follows from (4.226) that the stationary distribution of x is

$$w_{st}(x) = C \exp\left\{\frac{2}{\kappa}\int_{x_1}^{x} f(z)\,dz\right\}, \quad (4.227)$$

which is independent of μ and identical with the distribution of the one-dimensional Markov process described by the equation

$$\dot{x} = f(x) + \xi(t)\,.$$

In fact, this equation can be obtained either by setting $\mu = 0$ in (4.223) or by setting $g \equiv 1$ in (4.172). Therefore, the stationary distribution (4.227) is of course the same as (4.175) when $g \equiv 1$.

[9] In this equation and the next two, C denotes the normalization constant, which is suitably chosen in each case.

The stationary distributions (4.226) and (4.227) exist subject to the same restrictions on the function $f(x)$ as in the case of a one-dimensional Markov process. For a fluctuation equation of the form (4.224), the stationary solution (4.226) is replaced by

$$w_{st}(x, \dot{x}) = C \exp\left\{-\frac{ab}{\kappa} x^2 + \frac{2b}{\kappa}\left[\int_{x_1}^{x} f(z)\, dz + \langle g \rangle x\right]\right\}, \quad (4.228)$$

where

$$\kappa = \int_{-\infty}^{\infty} \mathbf{K}[g(\xi), g(\xi_\tau)]\, d\tau .$$

2. To find the nonstationary solution of equation (4.225), we write $w(x, y)$ in the form of an expansion

$$w(x, y) = \sum_{m,n=0}^{\infty} T_{mn}(t)\, X_m(x)\, Y_n(y) \quad (4.229)$$

in terms of the eigenfunctions $X_m(x)$, $Y_n(y)$ which satisfy the equations

$$\frac{\kappa}{2} \frac{\partial^2 X_m(x)}{\partial x^2} - \frac{\partial}{\partial x}[f(x)\, X_m(x)] + \lambda_m X_m(x) = 0 , \quad (4.230)$$

$$\frac{\kappa}{2} \frac{\partial^2 Y_n(y)}{\partial y^2} + \frac{\partial}{\partial y}[y Y_n(y)] + \lambda_n' Y_n(y) = 0 , \quad (4.231)$$

and the orthonormality conditions [cf. (4.56)]

$$\int X_m(x)\, X_k(x) \frac{dx}{X_0(x)} = \delta_{mk} , \qquad \int Y_n(y)\, Y_k(y) \frac{dy}{Y_0(y)} = \delta_{nk} .$$

Substituting (4.229) into (4.225), we obtain

$$\sum_{m,n} \mu^2 \dot{T}_{mn} X_m Y_n = -\mu \sum_{m,n} T_{mn}[X_m' y Y_n + f(x)\, X_m Y_n'] - \sum_{m,n} \lambda_n' T_{mn} X_m Y_n . \quad m$$

Using the relation

$$X_m' y Y_n + f(x)\, X_m Y_n' = \lambda_m \int X_m\, dx\, Y_n' - \lambda_n' X_m' \int Y_n\, dy ,$$

which follows from (4.230) and (4.231), we can transform the basic equation into

$$\sum_{m,n}(\mu^2\dot{T}_{mn} + \lambda_n' T_{mn})\, X_m Y_n = \mu \sum_{k,n} T_{kn}[\lambda_n' X_k' \int Y_n\, dy - \lambda_k \int X_k\, dx\, Y_n'] . \tag{4.232}$$

We now expand X_k' and $\int X_k dx$ in terms of the eigenfunctions X_m by writing

$$\begin{aligned} X_k'(x) &= \sum_m X_m(x)\, a_{mk} , \\ \int X_k(x)\, dx &= \sum_m X_m(x)\, b_{mk} , \end{aligned} \tag{4.233}$$

where the expansion coefficients

$$\begin{aligned} a_{mk} &= \int X_m X_k' \frac{dx}{X_0} , \\ b_{mk} &= \int X_m \left(\int X_k\, dx\right) \frac{dx}{X_0} \end{aligned} \tag{4.234}$$

obviously form matrices which are inverses of each other, i.e.,

$$\sum_k a_{mk} b_{kl} = \delta_{ml} .$$

We also introduce similar coefficients for the functions $Y_n'(y)$ and $\int Y_n(y)dy$. In fact, since equation (4.231) is the same as equation (4.68), it has the same eigenvalues $\lambda_n' = n$ and eigenfunctions

$$Y_n(y) = \sqrt{\frac{2}{\kappa}} \frac{1}{\sqrt{n!}} F^{(n+1)}\left(\sqrt{\frac{2}{\kappa}}\, y\right) \tag{4.235}$$

[cf. (4.69)]. It follows that

$$\begin{aligned} Y_n'(y) &= \sqrt{\frac{2}{\kappa}}\, \sqrt{n+1}\, Y_{n+1}(y) , \\ \int Y_n(y)\, dy &= \sqrt{\frac{\kappa}{2}} \frac{1}{\sqrt{n}}\, Y_{n-1}(y) . \end{aligned} \tag{4.236}$$

Substituting (4.233) and (4.236) into (4.232), replacing λ_n' by n, and

equating coefficients of identical functions $X_m Y_n$, we obtain the system of equations

$$\mu^2 \dot{T}_{mn} + nT_{mn} = \mu \sqrt{\frac{\kappa}{2}} \sqrt{n+1} \sum_k a_{mk} T_{k,n+1} - \mu \sqrt{\frac{2}{\kappa}} \sqrt{n} \sum_k b_{mk} \lambda_k T_{k,n-1} \quad (m, n = 0, 1, 2, \ldots), \tag{4.237}$$

which is equivalent to the two-dimensional Fokker-Planck equation.

When the parameter μ is sufficiently small, we can solve (4.237) by a perturbation method, which goes as follows: Setting $n = 0$ in (4.237), we obtain

$$\mu^2 p T_{m0} = \mu \sqrt{\frac{\kappa}{2}} \sum_k a_{mk} T_{k1},$$

where here (and from now on), p denotes the differentiation operator $\partial/\partial t$. But T_{k1} can be found by setting $n = 1$ in (4.237):

$$T_{k1} = \frac{\sqrt{2}\,\mu}{1+\mu^2 p} \sqrt{\frac{\kappa}{2}} \sum_l a_{kl} T_{l2} - \frac{\mu}{1+\mu^2 p} \sqrt{\frac{2}{\kappa}} \sum_l b_{kl} \lambda_l T_{l0}. \tag{4.238}$$

Then, to find T_{l2}, we set $n = 2$ in (4.237), obtaining

$$T_{l2} = \frac{\sqrt{3}\mu}{2+\mu^2 p} \sqrt{\frac{\kappa}{2}} \sum_j a_{lj} T_{j3} - \frac{\sqrt{2}\mu}{2+\mu^2 p} \sqrt{\frac{2}{\kappa}} \sum_j b_{lj} \lambda_j T_{j1},$$

which we then substitute in the right-hand side of (4.238). This procedure leads to the following expansion in even powers of the small parameter μ:

$$\mu^2 p T_{m0} = -\frac{\mu^2 \lambda_m}{1+\mu^2 p} T_{m0} + \frac{2\mu^4}{(1+\mu^2 p)^2 (2+\mu^2 p)} \sum_{k,l} a_{mk} \lambda_k b_{kl} \lambda_l T_{l0} + O(\mu^6). \tag{4.239}$$

After deciding upon a given accuracy, we can discard all terms of higher orders, while simplifying the remaining terms somewhat. Thus, in the second approximation, we have

$$\left[p + \frac{\lambda_m}{1 + \mu^2 p}\right] T_{m0} = \mu^2 \sum_{k,l} a_{mk}\lambda_k b_{kl}\lambda_l T_{l0} + O(\mu^4)\,.$$

Transforming the expression in brackets, we obtain

$$\begin{aligned} p + \lambda_m(1 - \mu^2 p)^{-1} &= p + \lambda_m - \mu^2\lambda_m p + O(\mu^4) \\ &= (1 - \mu^2\lambda_m)\,[p + \lambda_m\,(1 - \mu^2\lambda_m)^{-1}] + O(\mu^4) \\ &= (1 - \mu^2\lambda_m)\,[p + \lambda_m + \mu^2\lambda_m^2] + O(\mu^4)\,. \end{aligned}$$

Therefore, to the same accuracy as (4.239), we can write

$$\dot{T}_{m0} = -(\lambda_m + \mu^2\lambda_m^2)\,T_{m0} + \mu^2 \sum_{k,l} a_{mk}\lambda_k b_{kl}\lambda_l T_{l0} + O(\mu^4)\,. \qquad (4.240)$$

The equation (4.240) describes how the coefficients T_{m0} vary in time. If $\mu = 0$, each coefficient T_{m0} decreases exponentially with "time constant" λ_m^{-1}, just as in the case of one-dimensional Markov processes (see Sec 4). Thus, if $\mu = 0$, a perturbation which has the form of an eigenfunction $X_m(x)$ decays without changing its shape. However, if $\mu \neq 0$, the perturbation still falls off exponentially, but it has a somewhat different form. We now illustrate this situation by examining the form of the perturbation which falls off most slowly (with time dependence given by the factor $e^{p_1 t}$), and which coincides with the first eigenfunction $X_1(x)$ when $\mu = 0$. Setting $\partial/\partial t = p_1$ in (4.240), we find that

$$T_{m0} = \frac{\mu^2}{p_1 + \lambda_m + \mu^2\lambda_m^2} \sum_k a_{mk}\lambda_k b_{k1}\lambda_1 T_{10} + O(\mu^4)\,, \qquad (4.241)$$

for $m \neq 1$. Moreover, setting $m = 1$, we find that the constant p_1, which determines the rate of decay of the perturbation, is equal to

$$p_1 = -\lambda_1 - \mu^2\lambda_1^2 + \mu^2\lambda_1 \sum_k a_{mk}\lambda_k b_{k1} + O(\mu^4)\,. \qquad (4.242)$$

The one-dimensional probability density $w(x)$ is given by the formula

$$w(x) = \sum_{m=0}^{\infty} T_{m0} X_m(x) , \tag{4.243}$$

which can be obtained by integrating the expansion (4.229) with respect to y. Substituting (4.241) into (4.243), and retaining only terms of order μ^2, we find that the "nonstationary perturbation" which falls off most slowly while the stationary distribution is being established has the form

$$w(x) = T_{10} \left[X_1(x) + \mu^2 \lambda_1 \sum_{m \neq 1} \frac{a_{mk} \lambda_k b_{k1}}{\lambda_m - \lambda_1} X_m(x) \right] , \tag{4.244}$$

and is proportional to the "modified eigenfunction" corresponding to $X_1(x)$. Similarly, we can find the shapes and time constants of the other modified eigenfunctions.

★ It should be noted that equation (4.240) is equivalent to a certain differential equation for the one-dimensional probability density $w(x)$. To derive this equation, we multiply (4.240) by $X_m(x)$ and sum the result over m, as suggested by the form of (4.243). In doing this, we take into account the relations

$$\sum_{m,k} X_m(x)\, a_{mk} C_k = \frac{\partial}{\partial x} \sum_k X_k(x)\, C_k ,$$

$$\sum_m X_m(x)\, \lambda_m C_m = \left[\frac{\partial}{\partial x} f(x) - \frac{\kappa}{2} \frac{\partial^2}{\partial x^2} \right] \sum_m X_m(x)\, C_m ,$$

which follow from (4.233) and (4.230). Thus, the coefficients a_{mk} correspond to the differentiation operation, the coefficients b_{kl} correspond to the integration operation, and the λ_m correspond to the operation

$$\frac{\partial}{\partial x} \left[f(x) - \frac{\kappa}{2} \frac{\partial}{\partial x} \right] .$$

Then a simple calculation leads to the differential equation

$$\dot{w}(x) = \frac{\partial}{\partial x}\left[1 - \mu^2 f'(x)\right]\left[-f(x)\, w(x) + \frac{\kappa}{2}\frac{\partial w(x)}{\partial x}\right], \qquad (4.245)$$

where $f'(x) = \partial f(x)/\partial x$. If this equation is regarded as an ordinary one-dimensional Fokker-Planck equation, the corresponding functions $K_1(x)$ and $K_2(x)$ are just

$$K_1(x) = f(x) - \mu^2 f(x) f'(x) - \mu^2 \frac{\kappa}{2} f''(x),$$
$$K_2(x) = \kappa[1 - \mu^2 f'(x)].$$

Naturally, if higher-order terms in μ are taken into account, the corresponding equation will be more complicated.★

3. If the parameter η is not small, another method has to be used to find the solution of equation (4.223). Thus, we now consider the opposite extreme, i.e., the case where η is large. Writing $1/\mu = \epsilon$ and replacing t by t/ϵ, we can transform equation (4.223) into the new equation

$$\ddot{x} + \epsilon\dot{x} - f(x) = \sqrt{\epsilon}\,\xi(t). \qquad (4.246)$$

If ϵ is small, this equation describes the behavior of a system performing nonlinear oscillations under the influence of weak frictional forces and weak external fluctuations. It is convenient to introduce the "energy"

$$E = \tfrac{1}{2}\dot{x}^2 + u(x), \qquad (4.247)$$

where

$$u(x) = -\int_{x_1}^{x} f(z)\, dz$$

is the "potential function." Multiplying (4.246) by $\dot{x}$, we obtain

$$\dot{E} = -\epsilon\dot{x}^2 + \sqrt{\epsilon}\,\dot{x}\xi(t). \qquad (4.248)$$

Thus, from (4.247) and (4.248), we obtain two fluctuation equations

$$\dot{x} = \sqrt{2[E - u(x)]},$$
$$\dot{E} = -2\epsilon[E - u(x)] + \sqrt{2\epsilon[E - u(x)]}\,\xi(t). \qquad (4.249)$$

Then, in the usual way, we can associate (4.249) with the following Fokker-Planck equation for the probability density $w(E, x)$:

$$\dot{w}(x, E) = -\frac{\partial}{\partial x}\left[\sqrt{2[E-u)}\, w\right] + 2\epsilon \frac{\partial}{\partial E}\left[\left(E - u - \frac{\kappa}{4}\right) w\right] + \epsilon\kappa \frac{\partial^2}{\partial E^2}\left[(E-u)\, w\right], \tag{4.250}$$

where

$$\kappa = \int_{-\infty}^{\infty} \langle \xi \xi_\tau \rangle \, d\tau .$$

If ϵ is small, the energy E is conserved during a large number of periods of the oscillations, and moreover, the time which $x(t)$ spends at the point x is inversely proportional to the velocity $\dot{x} = \sqrt{2(E-u)}$. Therefore, for a fixed value of the energy, the conditional probability density is

$$w(x \mid E) = \begin{cases} \text{const} \cdot [E - u(x)]^{-1/2} & \text{for } u(x) < E, \\ 0 & \text{for } u(x) > E. \end{cases} \tag{4.251}$$

Normalizing this distribution, we can write the two-dimensional probability density in the form

$$w(x, E) = w(E)\, w(x \mid E) = \frac{w(E)}{2\varphi'(E)\sqrt{E - u(x)}}, \tag{4.252}$$

Here

$$\varphi'(E) = \frac{1}{2}\int_{R(E)} \frac{dx}{\sqrt{E - u(x)}}, \tag{4.253}$$

and the integration is over the region $R(E)$ where $u(x) < E$.

The distribution (4.252) can also be derived directly from equation (4.250). In fact, after a "quasi-equilibrium distribution" has been established, characterized by the fact that w is of order ϵ, it follows from (4.250) that

$$\frac{\partial}{\partial x}\left[\sqrt{E - u(x)}\, w\right] = O(\epsilon),$$

which when integrated gives

$$w = \frac{c(E)}{\sqrt{E - u(x)}} + O(\epsilon).$$

Thus, the deviation of $w(E, x)$ from (4.252) is a quantity whose order of magnitude is ϵ.

Substituting (4.252) into (4.250) and integrating with respect to x, we obtain the one-dimensional Fokker-Planck equation

$$\dot{w}(E) = \epsilon \frac{\partial}{\partial E}\left[\left(\frac{\varphi(E)}{\varphi'(E)} - \frac{\kappa}{2}\right) w\right] + \epsilon \frac{\kappa}{2} \frac{\partial^2}{\partial E^2}\left[\frac{\varphi(E)}{\varphi'(E)} w\right] \quad (4.254)$$

for the energy, where

$$\varphi(E) = \int_{R(E)} \sqrt{E - u(x)}\, dx$$

[cf. (4.253)]. This reduction of the two-dimensional Fokker-Planck equation to a one-dimensional equation considerably simplifies the problem of investigating nonstationary noise problems. Of course, if we take account of (4.252) and (4.247), the stationary solution of (4.254) reduces to the solution (4.228) found previously.

11.2. Next, we consider the somewhat different equation

$$\ddot{x} + \epsilon h(x, \dot{x}) - f(x) = \sqrt{\epsilon}\, \xi(t). \quad (4.255)$$

In this case, the stationary solution of the corresponding Fokker-Planck equation can no longer be written so simply in terms of quadratures, as in the case of equation (4.224). However, for small values of the parameter ϵ, we can apply the method just described, which, in particular, allows us to find the stationary distribution.

Thus, multiplying (4.255) by $\dot{x}$ and introducing the energy (4.247) instead of $\dot{x}$, we obtain the equations

$$\begin{aligned} \dot{x} &= \sqrt{2(E - u)}, \\ \dot{E} &= -\epsilon \sqrt{2(E - u)}\, h[x, \sqrt{2(E - u)}] + \sqrt{2\epsilon(E - u)}\, \xi(t). \end{aligned} \quad (4.256)$$

The Fokker-Planck equation corresponding to (4.256) differs from (4.250) only by the form of the second term in the right-hand side, where instead of $2(E - u)$, we now have

$$\sqrt{2(E-u)}\, h[x, \sqrt{2(E-u)}]\,.$$

Substituting (4.252) into this Fokker-Planck equation and integrating with respect to x, we obtain the one-dimensional equation

$$\dot{w}(E) = \epsilon \frac{\partial}{\partial E}\left[\left(\frac{\psi(E)}{\varphi'(E)} - \frac{\kappa}{2}\right) w\right] + \epsilon \frac{\kappa}{2} \frac{\partial^2}{\partial E^2}\left[\frac{\varphi(E)}{\varphi'(E)} w\right], \quad (4.257)$$

where

$$\psi(E) = \frac{1}{\sqrt{2}} \int_{R(E)} h[x, \sqrt{2(E-u)}]\, dx\,. \quad (4.258)$$

From (4.258), we can find the stationary distribution in the usual way:

$$w_{st}(E) = \text{const} \cdot \varphi'(E) \exp\left\{-\frac{2}{\kappa} \int_{E_1}^{E} \frac{\psi(E')}{\varphi(E')}\, dE'\right\}. \quad (4.259)$$

Using the relation

$$\frac{d}{dE} F(E) = \frac{\psi(E)}{\varphi(E)} \quad (4.260)$$

to introduce the function $F(E)$, we can write the two-dimensional distribution (4.252) in the form

$$w_{st}(x, E) = \frac{\text{const}}{2\sqrt{E - u(x)}} \exp\left\{-\frac{2}{\kappa} F(E)\right\}. \quad (4.261)$$

Going over to $\dot{x}$ and bearing in mind that $dx\ dE = \dot{x}\, d\dot{x}\, dx$, we have

$$w_{st}(x, \dot{x}) = \frac{1}{N} \exp\left\{-\frac{2}{\kappa} F\left[\frac{1}{2}\dot{x}^2 + u(x)\right]\right\}, \quad (4.262)$$

where

$$N = \int \exp\left\{-\frac{2}{\kappa} F\left[\frac{1}{2}y^2 + u(x)\right]\right\} dx\, dy\,.$$

If we integrate (4.262) with respect to $\dot{x}$, we find that the stationary distribution of the coordinate x is

$$w_{st}(x) = \frac{1}{N}\int \exp\left\{-\frac{2}{\kappa}F\left[\frac{1}{2}y^2 + u(x)\right]\right\} dy\,. \tag{4.263}$$

When $h(x, \dot{x}) = \beta\dot{x}$, the functions $\psi(E)$ and $\varphi(E)$ are proportional, and then the distribution (4.262) coincides with (4.228).

11.3. Finally, we consider the equation

$$\mu^3\ddot{x} + [1 + \mu^3 h(x)]\,\dot{x} - f(x) = \xi(t)\,, \tag{4.264}$$

which is more general than (4.223). The corresponding Fokker-Planck equation is

$$\begin{aligned}\mu^2\dot{w}(x, y) = &-\mu\left[y\frac{dw}{\partial x} + f(x)\frac{\partial w}{\partial y}\right]\\ &+ \mu^2 h(x)\frac{\partial}{\partial y}(yw) + \frac{\partial}{\partial y}(yw) + \frac{\kappa}{2}\frac{\partial^2 w}{\partial y^2}\,,\end{aligned} \tag{4.265}$$

where $y = \mu\dot{x}$. We shall assume that the parameter μ is small, exploiting this fact to solve (4.265) by a method resembling that used on pp. 110-114.

First, we write the probability density $w(x, y)$ in the form

$$w(x, y) = \sum_n w_n(x)\, Y_n(y)\,, \tag{4.266}$$

where the $Y_n(y)$ are the eigenfunctions (4.235) of the equation (4.231). Next, we substitute this expansion into (4.265), using the first of the equations (4.236) and the relations

$$\begin{aligned}\frac{\partial}{\partial y}(yY_n) &= -\sqrt{n+1}\,\sqrt{n+2}\,Y_{n+2} - nY_n\,,\\ yY_n &= -\sqrt{\frac{\kappa}{2}}\left[\sqrt{n+1}\,Y_{n+1} + \sqrt{n}\,Y_{n-1}\right],\end{aligned} \tag{4.267}$$

which follow from (4.231) and (4.236). Then, grouping coefficients

of identical eigenfunctions in (4.265), we obtain the system of equations

$$\begin{aligned}(n+\mu^2 p+\mu^3 nh)\,w_n &= \mu\sqrt{\frac{\kappa}{2}}\sqrt{n+1}\,\frac{\partial w_{n+1}}{\partial x}\\ &+\mu\sqrt{\frac{2}{\kappa}}\sqrt{n}\left(-fw_{n-1}+\frac{\kappa}{2}\frac{\partial w_{n-1}}{\partial x}\right.\\ &\left.-\mu^2\sqrt{n-1}\sqrt{n}\,hw_{n-2}\right),\end{aligned}\tag{4.268}$$

where $p = \partial/\partial t$. This system can be solved by the method of successive approximations. If we choose the same accuracy as on pp. 112-113, then it is sufficient to retain only the first three equations, which correspond to $n = 0, 1, 2$ and have the form

$$\begin{aligned}\mu^2 p w_0 &= \mu\sqrt{\frac{\kappa}{2}}\frac{\partial}{\partial x}w_1,\\ w_1 &= \mu\sqrt{\frac{\kappa}{2}}\frac{\sqrt{2}}{1+\mu^2 p+\mu^2 h}\frac{\partial}{\partial x}w_2+\mu\sqrt{\frac{2}{\kappa}}\frac{1}{1+\mu^2 p+\mu^2 h}Dw_0,\\ w_2 &= \mu\sqrt{\frac{\kappa}{2}}\frac{\sqrt{3}}{2+\mu^2 p+2\mu^2 h}\frac{\partial}{dx}w_3\\ &+\mu\sqrt{\frac{2}{\kappa}}\frac{\sqrt{2}}{2+\mu^2 p+2\mu^2 h}Dw_1-\frac{\mu^2\sqrt{2}\,h}{2+\mu^2 p+2\mu^2 h}w_0,\end{aligned}\tag{4.269}$$

where

$$D = -f+\frac{\kappa}{2}\frac{\partial}{\partial x}.$$

Substituting the second of these equations into the first and third, and using the third equation as well, we obtain an equation which, to the required degree of accuracy, can be written as

$$pw_0 = \frac{\partial}{\partial x}\frac{1}{1+\mu^2 p+\mu^2 h}Dw_0+\mu^2\frac{\partial^2}{\partial x^2}D^2 w_0-\mu^2\frac{\kappa}{2}\frac{\partial^2}{\partial x^2}hw_0+O(\mu^4).$$

Transforming $(1+\mu^2 p+\mu^2 h)^{-1}$ into $1-\mu^2 p-\mu^2 h$, transposing

the term $-\mu^2 p(\partial/\partial x)\, Dw_0$ to the left-hand side, and dividing both sides of the equation by $1 + \mu^2(\partial/\partial x)\, D$, we obtain

$$pw_0 = \frac{\partial}{\partial x} Dw_0 - \mu^2 \left(\frac{\partial}{\partial x} D\right)^2 w_0 + \mu^2 \frac{\partial^2}{\partial x^2} D^2 w_0 - \mu^2 \frac{\partial}{\partial x} hDw_0 - \mu^2 \frac{\kappa}{2} \frac{\partial^2}{\partial x^2} hw_0 . \tag{4.270}$$

Then, using the form of the operator D, we combine the second and third terms in the right-hand side of (4.270):

$$\frac{\partial^2}{\partial x^2} D^2 - \left(\frac{\partial}{\partial x} D\right)^2 = \frac{\partial}{\partial x}\left[\frac{\partial}{\partial x} D - D \frac{\partial}{\partial x}\right] D = -\frac{\partial}{\partial x}\left[\frac{\partial}{\partial x} f - f \frac{\partial}{\partial x}\right] D = -\frac{\partial}{\partial x} f' D .$$

Thus, equation (4.270) can be written in the form

$$\dot{w}(x) = \frac{\partial}{\partial x}\left\{[1 - \mu^2 f' - \mu^2 h]\left[-fw + \frac{\kappa}{2}\frac{\partial w}{\partial x}\right]\right\} - \frac{\kappa}{2}\frac{\partial^2}{\partial x^2}(\mu^2 hw) . \tag{4.271}$$

Here, we have replaced $w_0(x)$ by $w(x)$, since $w_0(x)$ is just the one-dimensional probability density of the coordinate x, as can be verified by integrating the expansion (4.266) with respect to y. Thus, we have obtained a one-dimensional Fokker-Planck equation which is equivalent to the two-dimensional equation (4.265), to the degree of accuracy chosen. In particular, using (4.271), we can find the stationary distribution

$$w_{st}(x) = \text{const} \cdot \exp\left\{\frac{2}{\kappa}\int_{x_1}^{x} f(z)\,[1 + \mu^2 h(z)]\, dz + \mu^2 h(x)\right\}, \tag{4.272}$$

which is a generalization of formula (4.227). If the fluctuation equation has the form

$$a\ddot{x} + b(x)\,\dot{x} = f(x) + \xi(t) , \tag{4.273}$$

then (4.272) becomes

$$w_{st}(x) = \text{const} \cdot \exp\left\{\frac{2}{\kappa}\int_{x_1}^{x} f(z)\, b(z)\, dz + b(x)\right\}. \tag{4.274}$$

12. Transition from a Markov Process to a Smoothed Process

The study of Markov processes is particularly appropriate, since effective mathematical methods are available for analyzing them. However, a certain care must be taken in replacing an actual process by a Markov process, since Markov processes have many special features, and, in particular, differ from the processes encountered in radio engineering by their lack of smoothness.

For example, suppose that the process $y(t)$ is described by the equation

$$\dot{y} - f(y) = \xi(t). \tag{4.275}$$

If the process $\xi(t)$ has a small correlation time $\tau_{cor} \ll \tau_0$, where, according to (4.183),

$$\tau_0 \sim \left(\frac{\partial f}{\partial y}\right)^{-1},$$

then, with the accuracy discussed in Sec. 8, the probability density $w(y)$ satisfies the Fokker-Planck equation (4.162). By writing (4.162), we replace the actual process $y(t)$ by a Markov process $x(t)$ which satisfies the equation

$$\dot{x} - f(x) = \zeta(t), \tag{4.276}$$

where $\zeta(t)$ is a delta-correlated random process with the same intensity K as $\xi(t)$. Since $\zeta(t)$ has infinite variance, it follows from (4.276) that the derivative $\dot{x}$ has infinite variance (i.e., $\mathbf{D}\dot{x} = \infty$), which indicates that the function $x(t)$ lacks smoothness and that its graph is "jagged." However, the inductances and capacitances always present in the circuits used in radio engineering act to smooth the process $x(t)$, so that to obtain a process like those encountered in practice, we have to subject $x(t)$ or $\zeta(t)$ to smoothing with a time constant $\tau_{sm} \ll \tau_0$. [In the case described by (4.275), τ_{sm} is of the order of the correlation time of the process $\xi(t)$.] The smoothing operation considerably modifies the behavior of the process $x(t)$ over time periods of the order $\Delta t \sim \tau_{sm}$, but the "large-scale" fluctuations (for which $\Delta t \sim \tau_0 \gg \tau_{sm}$) remain

almost unchanged. Thus, the results obtained by applying the techniques of Markov process theory are valuable only to the extent to which they characterize just these "large-scale" fluctuations.

To get an idea of how the operation of smoothing affects a Markov process, we consider a specific example of smoothing, obtained by setting

$$\xi(t) = \int_{-\infty}^{t} \frac{1}{\tau_{sm}} e^{-(t-t')/\tau_{sm}} \zeta(t')\, dt' . \tag{4.277}$$

This corresponds to assuming that $\xi(t)$ in (4.275) is an exponentially correlated random process, satisfying the equation

$$\tau_{sm}\dot{\xi} + \xi = \zeta(t) . \tag{4.278}$$

According to (4.275) and (4.278), we have the following second-order differential equation for the smoothed process $y(t)$:

$$\tau_{sm}\ddot{y} + [1 - \tau_{sm}f'(y)]\dot{y} - f(y) = \zeta(t) . \tag{4.279}$$

We arrive at approximately the same equation if in (4.276) we average $x(t)$ instead of $\zeta(t)$, by setting

$$y(t) = \int_{-\infty}^{t} \frac{1}{\tau_{sm}} e^{-(t-t')/\tau_{sm}} x(t')\, dt' , \tag{4.280}$$

i.e.,

$$\tau_{sm}\dot{y} + y = x . \tag{4.281}$$

Then (4.276) and (4.280) imply the equation

$$\tau_{sm}\ddot{y} + \dot{y} - f(\tau_{sm}\dot{y} + y) = \zeta(t) ,$$

which coincides with (4.279) if we expand $f(\tau_{sm}\dot{y} + y)$ in a power series in τ_{sm} and neglect terms of second and higher orders in τ_{sn}.

Equation (4.279) corresponds to a two-dimensional Markov process, and is a special case of equation (4.264), whose solution was studied in Sec. 11. Applying formula (4.272), we find that the stationary solution of (4.279) is

$$w_{st}(y) = \text{const} \cdot \exp\left\{\frac{2}{K}\int_{y_1}^{y} f(z)\, dz - \frac{\tau_{sm}}{K}f^2(y) - \tau_{sm}f'(y)\right\} . \tag{4.282}$$

Equation (4.282) shows the extent to which the smoothing perturbs the original stationary distribution.

If f is a linear function, i.e., if $f(x) = -\beta x$, then $x(t)$ is an exponentially correlated process, and

$$S[x; \omega] = \frac{S[\zeta; \omega]}{\omega^2 + \beta^2} = \frac{2K}{\omega^2 + \beta^2},$$

$$k_x(\tau) = \frac{K}{2\beta} e^{-\beta|\tau|}$$

(cf. entry 2 of Table 1, p. 25). The smoothed process $y(t)$ has the spectral density

$$S[y; \omega] = \frac{S[x; \omega]}{\omega^2 \tau_{sm}^2 + 1} = \frac{2K}{(\omega^2 \tau_{sm}^2 + 1)(\omega^2 + \beta^2)}. \qquad (4.283)$$

According to entry 7 of Table 1, $y(t)$ has the correlation function

$$k_y(\tau) = \frac{K}{2\beta}(1 - \beta^2 \tau_{sm}^2)^{-1} [e^{-\beta|\tau|} - \beta \tau_{sm} e^{-|\tau|/\tau_{sm}}], \qquad (4.284)$$

where

$$\mathrm{D}y = \frac{K}{2\beta}(1 + \beta \tau_{sm})^{-1}.$$

If $\beta \tau_{sm} \ll 1$, we can simplify (4.284) somewhat:

$$k_y(\tau) = \frac{K}{2\beta}[e^{-\beta|\tau|} - \beta \tau_{sm} e^{-|\tau|/\tau_{sm}}].$$

The smoothing operation just described is incomplete. In fact, (4.279) shows that the variance of the second derivative of $y(t)$ is infinite. To eliminate this infinite variance, while remaining within the framework of Markov process theory, we would have to consider a fluctuation equation of at least the third order. However, any random process actually encountered in radio engineering is analytic, and all its derivatives are finite with probability 1. Therefore, we cannot describe an actual process within the framework of Markov process theory, and the more accurately we wish to

approximate such a process by a Markov process, the more components the latter must have and the higher the order of the corresponding fluctuation equation must be.

As an example, we consider the analytic random process which has the correlation function

$$R(\tau) = \sigma^2 e^{-\beta^2\tau^2/4}$$

and the spectral density

$$S[\xi; \omega] = \frac{4\sigma^2}{\beta}\sqrt{\pi}\, e^{-\omega^2/\beta^2}. \tag{4.285}$$

(cf. entry 3 of Table 1). Equation (4.285) can be rewritten in the form

$$e^{\omega^2/\beta^2} S[\xi; \omega] = \frac{4\sigma^2}{\beta}\sqrt{\pi}. \tag{4.286}$$

If in (4.286) we make a power series expansion of e^{ω^2/β^2} and retain only the first $n + 1$ terms, then the process $\xi(t)$ is replaced by a Markov process $\xi_n(t)$ and (4.286) becomes

$$\left(1 + \frac{\omega^2}{\beta^2} + \ldots + \frac{1}{n!}\frac{\omega^{2n}}{\beta^{2n}}\right) S[\xi_n; \omega] = \frac{4\sigma^2}{\beta}\sqrt{\pi}. \tag{4.287}$$

In fact, we can represent the first factor in the left-hand side of (4.287) in the form $P_n(i\omega)P_n(-i\omega)$, where $P_n(p)$ is a suitable polynomial, and we can represent the right-hand side of (4.287) as the spectral density of a delta-correlated process $\zeta(t)$ by writing

$$S[\zeta; \omega] = \frac{4\sigma^2}{\beta}\sqrt{\pi}.$$

Then (4.286) becomes

$$|\, P_n(i\omega)\,|^2\, S[\xi_n; \omega] = S[\zeta; \omega],$$

which is equivalent to the following differential equation of order n:

$$P_n\left(\frac{d}{dt}\right)\xi_n = \zeta(t). \tag{4.288}$$

This implies that

$$\xi_n, \frac{d\xi_n}{dt}, \ldots, \frac{d^{n-1}\xi_n}{dt^{n-1}}$$

(i.e., the process $\xi_n(t)$ and its first $n-1$ derivatives) form the components of an n-dimensional Markov process. Moreover, the larger n, the less $\xi_n(t)$ differs from $\xi(t)$. If $n = 1$, the process $\xi_1(t)$ is an exponentially correlated Markov process satisfying the equation

$$\frac{1}{\beta}\dot{\xi}_1 + \xi_1 = \zeta(t),$$

and

$$\langle \xi_1 \xi_{1\tau} \rangle = \sigma^2 \sqrt{\pi}\, e^{-\beta|\tau|}.$$

The process $\xi_2(t)$, which is a closer approximation to $\xi(t)$, has the correlation function

$$\langle \xi_2 \xi_{2\tau} \rangle = \sigma^2 \sqrt{\frac{\pi}{\sqrt{2}+1}}\, e^{-a|\tau|}\left[\cos b\tau + \frac{a}{b}\sin b\,|\,\tau\,|\right], \quad (4.289)$$

where

$$a = \beta\sqrt{\frac{\sqrt{2}+1}{2}}, \qquad b = \beta\sqrt{\frac{\sqrt{2}-1}{2}}, \qquad \frac{a}{b} = \sqrt{2}+1,$$

(cf. entry 5 of Table 1), and satisfies the differential equation

$$\frac{1}{\beta^2\sqrt{2}}\ddot{\xi} + \frac{\sqrt{\sqrt{2}+1}}{\beta}\dot{\xi} + \xi = \zeta.$$

Supplement

★ Since the calculations leading to the formulas (4.179) are rather complicated, we give some of the details here: It follows from (4.151) and (4.145) that

$$\dot{L}(1 - L + L^2 - \ldots) = \epsilon A + \epsilon^2 B + \epsilon^3 C + \ldots,$$

where

$$C = -\frac{\partial}{\partial x}\langle \dot{H}_3\rangle + \frac{\partial^2}{\partial x^2}[\langle \dot{H}_2 H_1\rangle + \langle \dot{H}_1 H_2\rangle] - \frac{1}{2}\frac{\partial^3}{\partial x^3}\langle \dot{H}_1 H_1^2\rangle$$
$$-\frac{\partial}{\partial x}\Big[\langle \dot{H}_2\rangle \frac{\partial}{\partial x}\langle H_1\rangle + \langle \dot{H}_1\rangle \frac{\partial}{\partial x}\langle H_2\rangle\Big] + \frac{1}{2}\frac{\partial}{\partial x}\langle \dot{H}_1\rangle \frac{\partial^2}{\partial x^2}\langle H_1^2\rangle$$
$$+\frac{\partial^2}{\partial x^2}\langle \dot{H}_1 H_1\rangle \frac{\partial}{\partial x}\langle H_1\rangle - \frac{\partial}{\partial x}\langle \dot{H}_1\rangle \frac{\partial}{\partial x}\langle H_1\rangle \frac{\partial}{\partial x}\langle H_1\rangle . \qquad (4.290)$$

We write (4.290) in the form

$$C = -\frac{\partial}{\partial x}C_1 + \frac{\partial^2}{\partial x^2}C_2 - \frac{\partial^3}{\partial x^3}C_3 , \qquad (4.291)$$

where C_1, C_2 and C_3 are functions which can be expressed in terms of derivatives and integrals of the function F appearing in (4.142). If we use the relations

$$\frac{\partial}{\partial x}\langle \dot{H}_2\rangle \frac{\partial}{\partial x}\langle H_1\rangle = \frac{\partial^2}{\partial x^2}\langle \dot{H}_2\rangle \langle H_1\rangle - \frac{\partial}{\partial x}\Big\langle \frac{\partial \dot{H}_2}{\partial x}\Big\rangle \langle H_1\rangle ,$$

$$\frac{\partial}{\partial x}\langle \dot{H}_1\rangle \frac{\partial}{\partial x}\langle H_2\rangle = \frac{\partial^2}{\partial x^2}\langle \dot{H}_1\rangle \langle H_2\rangle - \frac{\partial}{\partial x}\Big\langle \frac{\partial \dot{H}_1}{\partial x}\Big\rangle \langle H_2\rangle ,$$

$$\frac{\partial}{\partial x}\langle \dot{H}_1\rangle \frac{\partial^2}{\partial x^2}\langle H_1^2\rangle = \frac{\partial^3}{\partial x^3}\langle \dot{H}_1\rangle \langle H_1^2\rangle$$
$$-2\frac{\partial^2}{\partial x^2}\Big\langle \frac{\partial \dot{H}_1}{\partial x}\Big\rangle \langle H_1^2\rangle + \frac{\partial}{\partial x}\Big\langle \frac{\partial^2 \dot{H}_1}{\partial x^2}\Big\rangle \langle H_1^2\rangle ,$$

$$\frac{\partial^2}{\partial x^2}\langle \dot{H}_1 H_1\rangle \frac{\partial}{\partial x}\langle H_1\rangle = \frac{\partial^3}{\partial x^3}\langle \dot{H}_1 H_1\rangle \langle H_1\rangle$$
$$-\frac{\partial^2}{\partial x^2}\Big\langle \frac{\partial \dot{H}_1}{\partial x} H_1\Big\rangle \langle H_1\rangle - \frac{\partial^2}{\partial x^2}\Big\langle \dot{H}_1 \frac{\partial H_1}{\partial x}\Big\rangle \langle H_1\rangle ,$$

$$\frac{\partial}{\partial x}\langle \dot{H}_1\rangle \frac{\partial}{\partial x}\langle H_1\rangle \frac{\partial}{\partial x}\langle H_1\rangle = \frac{\partial^3}{\partial x^3}\langle \dot{H}_1\rangle \langle H_1\rangle^2$$
$$-\frac{\partial^2}{\partial x^2}\Big[2\Big\langle \frac{\partial \dot{H}_1}{\partial x}\Big\rangle \langle H_1\rangle^2 + \langle \dot{H}_1\rangle \Big\langle \frac{\partial H_1}{\partial x}\Big\rangle \langle H_1\rangle\Big]$$
$$+\frac{\partial}{\partial x}\Big[\Big\langle \frac{\partial^2 \dot{H}_1}{\partial x^2}\Big\rangle \langle H_1\rangle^2 + \Big\langle \frac{\partial \dot{H}_1}{\partial x}\Big\rangle \Big\langle \frac{\partial H_1}{\partial x}\Big\rangle \langle H_1\rangle\Big] ,$$

which are consequences of the operator identity

$$f(x)\frac{\partial}{\partial x} = \frac{\partial}{\partial x}f(x) - \frac{\partial f(x)}{\partial x}.$$

Substituting these relations into (4.290), we obtain

$$C_1 = \langle \dot{H}_3 \rangle - \left\langle \frac{\partial \dot{H}_2}{\partial x} \right\rangle \langle H_1 \rangle - \left\langle \frac{\partial \dot{H}_1}{\partial x} \right\rangle \langle H_2 \rangle - \frac{1}{2}\left\langle \frac{\partial^2 \dot{H}_1}{\partial x^2} \right\rangle \langle H_1^2 \rangle + \left\langle \frac{\partial^2 \dot{H}_1}{\partial x^2} \right\rangle \langle H_1 \rangle^2 + \left\langle \frac{\partial \dot{H}_1}{\partial x} \right\rangle \left\langle \frac{\partial H_1}{\partial x} \right\rangle \langle H_1 \rangle,$$

$$C_2 = \langle \dot{H}_2 H_1 \rangle + \langle \dot{H}_1 H_2 \rangle - \langle \dot{H}_2 \rangle \langle H_1 \rangle - \langle \dot{H}_1 \rangle \langle H_2 \rangle - \left\langle \frac{\partial \dot{H}_1}{\partial x} \right\rangle \langle H_1^2 \rangle - \left\langle \frac{\partial \dot{H}_1}{\partial x} H_1 \right\rangle \langle H_1 \rangle - \left\langle \dot{H}_1 \frac{\partial H_1}{\partial x} \right\rangle \langle H_1 \rangle + 2\left\langle \frac{\partial \dot{H}_1}{\partial x} \right\rangle \langle H_1 \rangle^2 + \langle \dot{H}_1 \rangle \left\langle \frac{\partial H_1}{\partial x} \right\rangle \langle H_1 \rangle, \tag{4.292}$$

$$C_3 = \frac{1}{2}\langle \dot{H}_1 H_1^2 \rangle - \frac{1}{2}\langle \dot{H}_1 \rangle \langle H_1^2 \rangle - \langle \dot{H}_1 H_1 \rangle \langle H_1 \rangle + \langle \dot{H}_1 \rangle \langle H_1 \rangle^2.$$

Next, we substitute (4.146) and (4.147) into (4.292). Using (1.47), we easily see that

$$C_3 = \frac{1}{2}\mathbf{K}[\dot{H}_1, H_1, H_1] = \frac{1}{2}\int_{-\infty}^{0}\int_{-\infty}^{0} \mathbf{K}[F, F_\tau, F_\sigma]\, d\tau\, d\sigma.$$

Similarly, the expression for C_1 is found to be

$$C_1 = \frac{1}{2}\mathbf{K}\left[\frac{\partial^2 F}{\partial x^2}, H_1, H_1\right] + \left\langle \frac{\partial F}{\partial x}\int_{t_0}^{t} dt' \frac{\partial F}{\partial x}(t') \int_{t_0}^{t'} F(t'')\, dt'' \right\rangle - \left\langle \frac{\partial F}{\partial x}\int_{t_0}^{t} \frac{\partial F}{\partial x}(t')\, dt' \right\rangle \left\langle \int_{t_0}^{t} F(t'')\, dt'' \right\rangle - \left\langle \frac{\partial F}{\partial x} \right\rangle \left\langle \int_{t_0}^{t} dt' \frac{\partial F}{\partial x}(t') \int_{t_0}^{t'} F(t'')\, dt'' \right\rangle + \left\langle \frac{\partial F}{\partial x} \right\rangle \left\langle \int_{t_0}^{t} \frac{\partial F}{\partial x}(t')\, dt' \right\rangle \left\langle \int_{t_0}^{t} F(t'')\, dt'' \right\rangle.$$

We now divide the region of integration $t' < t$, $t'' < t$ of the third and last terms into two parts, i.e., the region $t' > t''$ and the region $t' < t''$. Then we use the identity

$$\langle \xi\eta\zeta\rangle - \langle\xi\rangle\langle\eta\zeta\rangle - \langle\xi\eta\rangle\langle\zeta\rangle + \langle\xi\rangle\langle\eta\rangle\langle\zeta\rangle = \mathbf{K}[\xi,\eta,\zeta] + \langle\eta\rangle\mathbf{K}[\xi,\zeta] .$$

This leads to the formula

$$C_1 = \frac{1}{2}\mathbf{K}\left[\frac{\partial^2 F}{\partial x^2}, H_1, H_1\right] + \iint\limits_{0>\tau>\sigma}\left\{\mathbf{K}\left[\frac{\partial F}{\partial x}, \frac{\partial F_\tau}{\partial x}, F_\sigma\right] + \left\langle\frac{\partial F_\tau}{\partial x}\right\rangle \mathbf{K}\left[\frac{\partial F}{\partial x}, F_\sigma\right]\right\} d\sigma\, d\tau$$
$$- \iint\limits_{\tau<\sigma<0}\left\{\left\langle\frac{\partial F}{\partial x}\frac{\partial F_\tau}{\partial x}\right\rangle - \left\langle\frac{\partial F}{\partial x}\right\rangle\left\langle\frac{\partial F_\tau}{\partial x}\right\rangle\right\}\langle F_\sigma\rangle\, d\tau\, d\sigma .$$

To obtain the final result, we need only make the change of variables $\sigma \to \tau$, $\tau \to \sigma$ in the last integral. An expression for C_2 is found by the same kind of argument. The derivation of (4.179) is now straightforward. ★

CHAPTER 5

Nonstationary Random Processes

In radio engineering, one encounters many random processes other than the stationary processes considered in Chap. 2. In fact, as the subject of statistical radio engineering evolves both experimentally and theoretically, one has to deal increasingly often with the extensive class of nonstationary processes. In this chapter, we shall consider only a few very important types of nonstationary processes.

1. Processes with Slow Nonstationary Changes

This kind of nonstationary process is characterized by the fact that it behaves almost like a stationary process for long time intervals, which, however, do not too greatly exceed the correlation time. Although the mean value $m_1(t)$ and the correlation functions $k_2(t, t+\tau)$, $k_3(t, t+\tau_1, t+\tau_2)$, ... depend on the absolute time t (and not just on the time differences $\tau, \tau_1, \tau_2, \ldots$), they change only slightly during times of the order of the correlation time τ_{cor}. As applied to the mean m_1 and the correlation function k_2, this condition can be expressed analytically as follows:

$$\frac{\partial m_1}{\partial t}\tau_{cor} \ll m_1, \qquad \frac{\partial k_2}{\partial t}\tau_{cor} \ll k_2. \tag{5.1}$$

For such a process, we can define a spectral density

$$S[\xi; \omega, t] = 2\int_{-\infty}^{\infty} m_2(t, t+\tau)\, e^{i\omega\tau}\, d\tau, \tag{5.2}$$

which varies slowly with the time t. In the integrand of (5.2), we can replace $m_2(t, t+\tau)$ by $m_2(t-\tau, t)$ or by the symmetric expression $m_2(t-\frac{1}{2}\tau, t+\frac{1}{2}\tau)$. The spectral density (5.2) for a process with slow nonstationary changes plays the same role as the ordinary spectral density for a stationary process.

The actual processes encountered in radio engineering always undergo slow nonstationary changes due to lack of stability of parameters of the equipment. These changes are unimportant if the time $k_n/(\partial k_n/\partial t)$ which it takes for the nonstationarity to manifest itself is much greater than the relevant time constants of the given problem. If this is not the case, one has to take the nonstationarity into account.

2. Switched-On Processes

In electrical equipment, noise processes do not become "established" at once; in fact, the influence of the initial conditions is "felt" for some time after the equipment is switched on. As time goes by, the influence of the initial conditions "wears off," and the "switched-on" process asymptotically approaches a stationary process.

As an example, consider a stationary normal process $\xi(t)$, with mean value m and correlation function $\sigma^2 R(\tau)$. If the noise process $\xi(t)$ has a known deterministic value $\xi(0) = a$ at the time $t = 0$, then, according to formula (3.26), the process is subsequently described by the conditional mean value

$$m_1(t) = m + (a - m)R(t)$$

and the conditional correlation function

$$k_2(t, t+\tau) = \sigma^2[R(\tau) - R(t)\,R(t+\tau)]\,.$$

These quantities depend on t, and hence the process $\xi(t)$ is nonstationary. However, after a time considerably greater than the correlation time has elapsed (i.e., when $t \gg \tau_{cor}$, $t+\tau \gg \tau_{cor}$), the functions $R(t)$, $R(t+\tau)$ approach zero, and $m_1(t)$, $k_2(t, t+\tau)$

approach the quantities m, $\sigma^2 R(\tau)$ corresponding to a stationary process.

In many cases, a noise process in a piece of electrical equipment is described by a differential equation, e.g., a first-order differential equation

$$\dot{\xi} = f(\xi) + \zeta(t), \tag{5.3}$$

involving an external random perturbation $\zeta(t)$ which is stationary. The function f is usually such that as time increases, a stationary process is established in the system. However, in the beginning, a nonstationary transient process takes place, which can be studied by using the methods given in Chap. 4, Sec. 4.

3. Processes with Stationary Increments

When the function $f(\xi)$ in equation (5.3) takes certain forms, the noise process $\xi(t)$ never "turns into" a stationary process, no matter how much time t has elapsed since the initial time. In particular, this is the case for the equation

$$\dot{\xi} = \zeta(t), \tag{5.4}$$

which describes a *process with stationary increments.* Because of the stationarity of the function $\zeta(t)$, the increment

$$\xi(t_0 + T) - \xi(t_0) = \int_{t_0}^{t_0+T} \zeta(t)\, dt$$

during an interval of length T has the same statistical properties, regardless of which time t_0 is chosen. (In what follows, for brevity we set $t_0 = 0$ and $\xi(t_0) = 0$.) If we subtract from the integral

$$\xi(t) = \int_0^t \zeta(t')\, dt' \tag{5.5}$$

its own mean value, and then square and average the result, we obtain

$$\mathrm{D}\xi(t) = \int_0^t \int_0^t k_\zeta(t_1 - t_2)\, dt_1\, dt_2 = 2\int_0^t (t - \tau)\, k_\zeta(\tau)\, d\tau. \tag{5.6}$$

Differentiating (5.6) with respect to t, we obtain

$$\frac{d}{dt}\mathbf{D}\xi(t) = 2\int_0^t k_\zeta(\tau)\,d\tau\,. \tag{5.7}$$

Sometimes it is convenient to write (5.7) in the form

$$\frac{d}{dt}\mathbf{D}\xi(t) = 2\int_0^\infty k_\zeta(\tau)\,d\tau - 2\int_t^\infty k_\zeta(\tau)\,d\tau$$

or

$$\frac{d}{dt}\mathbf{D}\xi(t) = \tfrac{1}{2}\,S[\zeta - \langle\zeta\rangle; 0] - 2\int_t^\infty k_\zeta(\tau)\,d\tau\,. \tag{5.8}$$

To determine the variance $\mathbf{D}\xi(t)$ as a function of time by using (5.7) or (5.8), we have to integrate both sides of the equation with respect to t, using the initial condition

$$\mathbf{D}\xi(0) = 0\,. \tag{5.9}$$

Once we have found $\mathbf{D}\xi(t)$, we can easily calculate the correlation function $k_\xi(t_1, t_2)$. In fact, if we substitute

$$a = \xi(t_1) - \langle\xi(t_1)\rangle\,, \qquad b = \xi(t_2) - \langle\xi(t_2)\rangle$$

into the identity

$$ab = \tfrac{1}{2}\,[a^2 + b^2 - (a - b)^2]\,,$$

and then average the result, we obtain

$$k_\xi(t_1, t_2) = \tfrac{1}{2}\,\{\mathbf{D}\xi(t_1) + \mathbf{D}\xi(t_2) - \mathbf{D}[\xi(t_1) - \xi(t_2)]\}\,.$$

However, because of the stationarity of the increments, we have

$$\mathbf{D}[\xi(t_1) - \xi(t_2)] = \mathbf{D}\xi(|\,t_1 - t_2\,|)\,,$$

and hence

$$k_\xi(t_1, t_2) = \tfrac{1}{2}\,\{\mathbf{D}\xi(t_1) + \mathbf{D}\xi(t_2) - \mathbf{D}\xi(|\,t_1 - t_2\,|)\}\,. \tag{5.10}$$

We now consider some examples.

Example 1. Let $\zeta(t)$ be a delta-correlated normal process with zero mean value. Substituting the correlation function $k_\zeta(\tau) = K\delta(\tau)$ into (5.6), we obtain

$$\mathrm{D}\xi(t) = Kt\,. \tag{5.11}$$

Therefore, the distribution of the increment of ξ during time t is

$$w(\xi) = \frac{1}{\sqrt{2\pi Kt}} \exp\left\{-\frac{\xi^2}{2Kt}\right\}.$$

Because of the delta-function form of the correlation function $k_\zeta(\tau)$, increments of the process $\xi(t)$ during nonoverlapping time intervals are independent. Therefore, $\xi(t)$ is a Markov process with transition probability

$$p(\xi_\tau, \xi) = \frac{1}{\sqrt{2\pi K\tau}} \exp\left\{-\frac{(\xi_\tau - \xi)^2}{2K\tau}\right\}. \tag{5.12}$$

Sometimes this kind of normal process with independent increments is called a *Wiener process*. According to (5.10), the increment (5.5) has the correlation function

$$k_\xi(t_1, t_2) = \frac{K}{2}\left[t_1 + t_2 - |\,t_1 - t_2\,|\,\right]. \tag{5.13}$$

Example 2. Suppose now that the stationary process $\zeta(t)$ has the spectral density

$$S[\zeta;\omega] = a\omega^{1-\alpha}\,, \tag{5.14}$$

where

$$0 \leqslant \alpha < 1\,.$$

Then, applying the formula

$$\int_0^\infty e^{-i\tau\omega}\omega^{p-1}\,d\omega = (i\tau)^{-p}\Gamma(p)\,,$$

we find that

$$k_\zeta(\tau) = \frac{a}{2\pi}\,\mathrm{Re}\int_0^\infty e^{-i\tau\omega}\omega^{1-\alpha}\,d\omega = \frac{a}{2\pi}\cos\left[(2-\alpha)\frac{\pi}{2}\right]\Gamma(2-\alpha)\,\tau^{\alpha-2} \tag{5.15}$$

for $\tau > 0$, or

$$k_\zeta(\tau) = \frac{a}{4} \frac{\tau^{\alpha-2}}{\Gamma(\alpha-1)\sin\frac{1}{2}\pi\alpha}, \tag{5.16}$$

where we have used the familiar formula

$$\Gamma(1-x)\,\Gamma(x) = \frac{\pi}{\sin \pi x}$$

involving the gamma function.

When $\tau = 0$, the expression (5.16) is not valid, since it can easily be seen that $k_\zeta(\tau)$ takes arbitrarily large values near $\tau = 0$. In fact, including the value $\tau = 0$ in (5.16) is like including a delta-function term of the form $c\delta(\tau)$, where c is infinite; the resulting infinite variance $\mathrm{D}\zeta = k_\zeta(0)$ is of the same character as the infinite variance of delta-correlated noise, and is related to the presence of high-frequency spectral components in $S[\zeta; \omega]$. However, in actual processes there are practically no spectral components with frequencies exceeding a certain upper cutoff frequency ω_c. Hence, the variance of an actual process is finite, and has the order of magnitude

$$k_\zeta(0) \sim \frac{a}{4} \frac{\omega_c^{2-\alpha}}{\Gamma(\alpha-1)\sin\frac{1}{2}\pi\alpha}. \tag{5.17}$$

As $\alpha \to 1$, the process $\zeta(t)$ turns into delta-correlated noise, while for $\alpha = 0$, formula (5.15) gives

$$k_\zeta(\tau) = -\frac{a}{2\pi}\frac{1}{\tau^2} \qquad (\tau > 0). \tag{5.18}$$

Next, we calculate the variance of the integrated process $\xi(t)$, using formula (5.8). Taking into account (5.14) and (5.16), we find that

$$\frac{d}{dt}\mathrm{D}\xi(t) = \frac{a}{2} \frac{t^{\alpha-1}}{(\alpha-1)\,\Gamma(\alpha-1)\sin\frac{1}{2}\pi\alpha}. \tag{5.19}$$

Integrating (5.19) with the condition (5.9), we obtain

$$\mathrm{D}\xi(t) = \frac{a}{2} \frac{t^{\alpha}}{\Gamma(\alpha+1)\sin\frac{1}{2}\pi\alpha} \qquad (0 < \alpha < 1). \tag{5.20}$$

The case $\alpha = 0$ requires special treatment. According to (5.8) and (5.18), we have

$$\frac{d}{dt}\mathrm{D}\xi(t) = \frac{a}{\pi t}. \tag{5.21}$$

The integral of (5.21) is divergent, but formal integration gives

$$\mathrm{D}\xi(t) = \frac{a}{\pi}(\ln t - \ln 0). \tag{5.22}$$

The infinite value $\ln 0$ is due to the presence of high frequencies. If we introduce an upper cutoff frequency ω_c, then the integration of (5.21) should be carried out from $1/\omega_c$ to t, rather than from 0 to t, so that formula (5.22) is replaced by

$$\mathrm{D}\xi(t) = \frac{a}{\pi}\ln \omega_c t \qquad (\alpha = 0). \tag{5.23}$$

From a knowledge of the variance $\mathrm{D}\xi(t)$, we can find the correlation function of the increment of $\xi(t)$ during the time t. Using (5.10), we obtain

$$k_\xi(t_1, t_2) = \frac{a}{4}\frac{t_1^\alpha + t_2^\alpha - |t_1 - t_2|^\alpha}{\Gamma(\alpha + 1)\sin \frac{1}{2}\pi\alpha} \qquad (0 < \alpha < 1) \tag{5.24}$$

and

$$k_\xi(t_1, t_2) = \frac{a}{2\pi}\ln \frac{\omega_c t_1 t_2}{|t_1 - t_2|} \qquad (\alpha = 0). \tag{5.25}$$

If we now fix the value of the difference $t_2 - t_1 = \tau$ and let the time $t = \frac{1}{2}(t_1 + t_2)$ approach infinity, then, as can be seen from the expressions just obtained, the correlation function will not approach a finite limit, so that the "switched-on" process $\xi(t)$ never achieves a stationary regime. However, under the condition $\tau \ll t$, the expressions (5.24) and (5.25) can be simplified somewhat. By making Taylor's series expansions with respect to τ, we find that

$$\left(t + \frac{\tau}{2}\right)^\alpha + \left(t - \frac{\tau}{2}\right)^\alpha = 2t^\alpha + \frac{\alpha(\alpha - 1)}{4}t^\alpha\frac{\tau^2}{t^2} + \ldots \qquad (0 < \alpha < 1),$$

$$\ln\left(t + \frac{\tau}{2}\right) + \ln\left(t - \frac{\tau}{2}\right) = 2\ln t - \frac{\tau^2}{2t^2} + \ldots \qquad (\alpha = 0).$$

Therefore, with an error of order $\tau^{\alpha}(\tau/t)^{2-\alpha}$, we have

$$k_{\xi}\left(t-\frac{\tau}{2},\ t+\frac{\tau}{2}\right)=\frac{a}{4}\frac{2t^{\alpha}-\tau^{\alpha}}{\Gamma(\alpha+1)\sin\frac{1}{2}\pi\alpha}\qquad(0<\alpha<1),\tag{5.26}$$

$$k_{\xi}\left(t-\frac{\tau}{2},\ t+\frac{\tau}{2}\right)=\frac{a}{2\pi}\ln\frac{\omega_c t^2}{\tau}\qquad(\alpha=0).\tag{5.27}$$

If the integrated process $\xi(t)$ were stationary, then it would have the spectral density

$$S[\xi;\omega]=\frac{1}{\omega^2}S[\zeta;\omega]=\frac{a}{\omega^{1+\alpha}},\tag{5.28}$$

corresponding to the correlation function

$$\begin{aligned}k_{\xi}(\tau)&=C-\frac{a}{4}\frac{\tau^{\alpha}}{\Gamma(\alpha+1)\sin\frac{1}{2}\pi\alpha}\qquad(0<\alpha<1),\\ k_{\xi}(\tau)&=C-\frac{\alpha}{2\pi}\ln\tau\qquad(\alpha=0),\end{aligned}\tag{5.29}$$

where the constant C is arbitrary. It is true that for an ideal process such that the formula (5.28) holds over the whole frequency range right down to zero frequency, the constant C has to be infinite in order to satisfy the requirement that the correlation function $k_{\xi}(\tau)$ be positive definite. However, for actual stationary processes, (5.28) holds only for frequencies $\omega>\omega_l$ which are not too small, i.e., which exceed a lower cutoff frequency ω_l, while for smaller frequencies $\omega<\omega_l$, $S[\xi;\omega]$ approaches a finite limit of order $a/\omega_l^{1+\alpha}$. Then, the constant C in (5.29) is finite and depends on ω_l, and (5.29) can be written in the form

$$\begin{aligned}k_{\xi}(\tau)&=\frac{a}{4}\frac{\omega_l^{-\alpha}-\tau^{\alpha}}{\Gamma(\alpha+1)\sin\frac{1}{2}\pi\alpha}\qquad(0<\alpha<1),\\ k_{\xi}(\tau)&=\frac{\alpha}{2\pi}\ln\frac{\omega_c}{\omega_l^2\tau}\qquad(\alpha=0).\end{aligned}\tag{5.30}$$

These considerations show that for observation times $t<1/\omega_l$, a process $\xi(t)$ with the spectral density (5.28) can be regarded as a nonstationary process with the correlation function (5.26) or

(5.27), while for observation times $t > 1/\omega_1$, $\xi(t)$ should be regarded as a stationary process with the correlation function (5.30). The difference between these two cases is tantamount to a difference in the interpretation of the constant C in (5.29).

4. Periodic Nonstationary Processes

According to (2.1), the statistical characteristics of a stationary process are invariant under arbitrary time shifts (corresponding to replacing t by $t + a$, where a is arbitrary). On the other hand, the statistical characteristics of a *periodic nonstationary process* are only invariant under shifts by a multiple of a certain *period* T_0 (corresponding to replacing t by $t + a$, where now $a = nT_0$ and n is an integer). As a result, the probability densities, moment functions and correlation functions of a periodic nonstationary process $\xi(t)$ depend on the absolute time, as well as on the time differences, but the dependence on the absolute time is periodic. If we represent the process $\xi(t)$ as a trigonometric series, its mean value and (second-order) correlation function can be written in the form

$$m(t) = \sum_{n=-\infty}^{\infty} m_n e^{2\pi i n(t/T_0)} = m_0 + 2\,\mathrm{Re}\sum_{n=1}^{\infty} m_n e^{2\pi i n(t/T_0)} \tag{5.31}$$

$$k\left(t - \frac{\tau}{2},\ t + \frac{\tau}{2}\right) = k_0(\tau) + 2\,\mathrm{Re}\sum_{n=1}^{\infty} k_n(\tau) e^{2\pi i n(t/T_0)}, \tag{5.32}$$

where in general the coefficients m_n and $k_n(\tau)$ are complex $[m_{-n} = m_n^*,\ k_{-n}(\tau) = k_n^*(\tau)]$.

If we shift $\xi(t)$ by an amount a, where $0 < a < T_0$, we obtain a random function $\xi_a(t) = \xi\,(t + a)$, whose statistical properties are different from those of the original process $\xi(t)$. In fact, in the general case, we have

$$\langle \xi_a(t) \rangle = m(t + a) \neq m(t),$$

$$\langle \xi_a(t_1)\xi_a(t_2) \rangle = m_2(t_1 + a,\ t_2 + a) \neq m_2(t_1, t_2),$$

... .

Using the language of radio engineering, we can say that the processes $\xi(t)$ and $\xi_a(t)$ have different *phases*. In some engineering systems, the phase of a periodic process is very important, and such systems are said to be *coherent*. In other systems, which are said to be *incoherent*, the phase of the process is unimportant, since the system does not respond to the phase, and instead carries out a time average. When dealing with incoherent sytems, we can replace a periodic nonstationary process by a stationary process, corresponding to a random phase "spread."

To illustrate this situation, we find the mean value $\tilde{m}$ and the correlation function $\tilde{k}$ of a random process $\xi(t) = \xi_a(t)$, whose phase a is a random variable which is statistically independent of $\xi(t)$ and has the uniform distribution

$$w(a) = \begin{cases} \dfrac{1}{T_0} & \text{for } 0 < a < T_0, \\ 0 & \text{otherwise.} \end{cases}$$

Averaging first with respect to $\xi(t)$ and then with respect to a, we find that

$$\begin{aligned} \tilde{m} &= \langle m(t+a)\rangle = \frac{1}{T_0}\int_0^{T_0} m(t+a)\,da, \\ \tilde{m}_2(t_1, t_2) &= \frac{1}{T_0}\int_0^{T_0} m_2(t_1+a, t_2+a)\,da. \end{aligned} \tag{5.33}$$

Setting

$$m_2(t, t') = m(t)m(t') + k(t, t')$$

in (5.33), using the formulas (5.31) and (5.32), and then carrying out the indicated integrations, we find that

$$\tilde{m} = m_0,$$

$$\tilde{m}_2(t_1, t_2) = \sum_{n=-\infty}^{\infty} m_n m_{-n} e^{2\pi i n(t_1-t_2)/T_0} + k_0(t_1 - t_2).$$

Subtracting $\tilde{m}^2$ from $m_2(t_1, t_2)$, we find that the correlation function is

$$\tilde{k}(\tau) = 2\sum_{n=1}^{\infty} |m_n|^2 \cos\frac{2\pi n\tau}{T_0} + k_0(\tau). \tag{5.34}$$

To calculate the higher-order moment functions and the probability densities, we can carry out a similar phase averaging. In particular, the two-dimensional probability density is

$$\tilde{w}(\xi_1, \xi_2, t_2 - t_1) = \frac{1}{T_0}\int_0^{T_0} w_\xi(\xi_1, \xi_2, t_1 + a, t_2 + a)\, da\,. \quad (5.35)$$

As a result of this averaging, the only dependence which remains is that involving the time difference τ, so that the process $\xi(t)$ is in fact stationary.

The above formulas can also be applied to the case where the original process $\xi(t)$ is periodic, but not random. For example, suppose that

$$\xi(t) = A \sin \omega_0 t = \frac{A}{2i}\,(e^{i\omega_0 t} - e^{-i\omega_0 t})\,. \quad (5.36)$$

Then

$$T_0 = \frac{2\pi}{\omega_0}, \qquad m_1 = \frac{A}{2i}, \qquad m_0 = m_2 = \ldots = 0\,, \qquad k_2 = 0\,,$$

and formula (5.34) gives

$$\tilde{k}(\tau) = \frac{A^2}{2}\cos \omega_0 \tau\,.$$

The function (5.36) corresponds to the two-dimensional probability density

$$w_\xi(\xi, \xi', t, t') = \delta(\xi - A \sin \omega_0 t)\,\delta(\xi' - A \sin \omega_0 t')\,.$$

Substituting this expression into (5.35), we find the two-dimensional distribution of the stationary process $\tilde{\xi}(t) = A \sin(\omega_0 t + \varphi)$, where the random phase $\varphi = \omega_0 a$ has a uniform distribution:

$$\begin{aligned}\tilde{w}(\xi_1, \xi_2, \tau) &= [2\pi\sqrt{A^2 - \xi_1^2}]^{-1}[\delta(\xi_2 - \xi_1 \cos \omega_0 \tau - \sqrt{A^2 - \xi_1^2}\sin \omega_0 \tau) \\ &\quad + \delta(\xi_2 - \xi_1 \cos \omega_0 \tau + \sqrt{A^2 - \xi_1^2}\sin \omega_0 \tau] \qquad (5.37) \\ &= \frac{1}{\pi}\,|\sin \omega_0 \tau\,|\,\delta(\xi_1^2 + \xi_2^2 - 2\xi_1\xi_2 \cos \omega_0 \tau - A^2 \sin^2 \omega_0 \tau)\,.\end{aligned}$$

CHAPTER 6

Systems of Random Points and Related Random Functions

As radio communication, automatic control, telemetry and other fields of electrical engineering continue to evolve, extensive use is made of pulse signals, where the positions, durations and other pulse parameters are often random, with various kinds of correlation between different pulses. Thus, in order to design radio equipment in a rational way, we have to know the statistical characteristics of random processes which are due to the superposition of random pulses. In this chapter, we shall devote special attention to the problem of calculating the spectral density of pulse sequences. The theory of random points presented here can be applied not only to pulse signals, but also to other cases where a random process can be decomposed into discrete elements or events, e.g., in the statistical theory of many-particle interactions.[1]

1. Methods for Describing Systems of Random Points

In many cases, one has to deal with a set of points which have random positions in a given space, in particular, on the time axis. Such points will simply be called *random points*. The times at which

[1] For an application of the theory of random points to a problem of radar detection, see R. L. Stratonovich, *The conditional distribution of correlated random points and the use of correlations for optimum extraction of a pulse signal from noise*, Izv. Akad. Nauk. SSSR, Otd. Tekh. Nauk, Energ. i Avtomat., no. 2, 148 (1961).

various random events terminate, the points of intersection of a random function with a given curve, the centers of gravity of particles in statistical physics, etc., are all examples of random points. Of course, by specifying the coordinates of the random points, we might convert this problem into one pertaining to a family of random variables; however, this way of doing things is inconvenient, since the number of random points is often a random variable itself, and when this is the case, concepts involving space and time relationships (e.g., the concept of stationarity) lose their intuitive meaning.

In the special case where there are no correlations between the positions of the random points, the distribution of the points is described by the *Poisson law*. However, the theory presented in this chapter includes all possible kinds of correlation between the random points, and hence the formulas that we derive are valid in the general case. In what follows, we consider the case of identical random points in a one-dimensional space, which can be interpreted as the time axis, but all our concepts and formulas can be generalized to the case of systems of points in a multidimensional space and to the case of points of several kinds.

A system of random points is completely characterized by the sequence of *distribution functions*

$$f_1(t_1), \qquad f_2(t_1, t_2), \qquad f_3(t_1, t_2, t_3), \ldots, \tag{6.1}$$

where $t_1, t_2, t_3, \ldots$ lie in the *domain of definition* of the system of random points (i.e., the set of possible positions of the points). Here, distribution function $f_s(t_1, \ldots, t_s)$ gives the probability

$$dP = [f_s(t_1, \ldots, t_s) + o(\Delta)]\Delta_1 \ldots \Delta_s$$

that at least one random point falls in each of the intervals

$$t_1 \leqslant t \leqslant t_1 + \Delta_1, \ldots, t_s \leqslant t \leqslant t_s + \Delta_s \qquad (\Delta_1 \leqslant \Delta, \ldots, \Delta_s \leqslant \Delta). \tag{6.2}$$

It is assumed that the values $t_1, \ldots, t_s$ do not coincide ($t_i \neq t_j$ if $i \neq j$) and that Δ is so small that the domains (6.2) do not overlap. For random points characterized by continuous statistical

properties, the distribution functions (6.1) are continuous and the values which they take when arguments coincide, e.g., when $t_i = t_j$, are conveniently defined by the corresponding limits, e.g., as $t_i \to t_j$. A desirable feature of these distribution functions is that they characterize local statistical properties at the positions corresponding to their arguments, without reference to statistical properties at other points in space.

We now choose an interval $0 \leqslant t \leqslant T$ in the domain of definition of the system of random points, and we consider the *generating functional*

$$L_T[v(t)] = \left\langle \prod_{j=1}^{n} [1 + v(t_j)] \right\rangle, \tag{6.3}$$

where n is the number of random points lying in the interval $0 \leqslant t \leqslant T$, and $t_1, \ldots, t_n$ are their position coordinates. The class of functions $v(t)$ is restricted by certain requirements which depend on the nature of the random points and will not be discussed here. Then different systems of random points are characterized by different generating functionals, as illustrated by the following simple systems.

Example 1. Suppose we know that a fixed number N of random points falls in the interval $0 \leqslant t \leqslant T_1$, where $T_1 > T$. Moreover, let each point take any position in $0 \leqslant t \leqslant T_1$ with equal probability, and let this position be statistically independent of the positions of the other points. Then the position of each point is described by the probability density

$$w(t_j) = \frac{1}{T_1} \text{ for } 0 \leqslant t \leqslant T_1 \qquad (1 \leqslant j \leqslant N),$$

and therefore

$$\langle 1 + v(t_j) \rangle = 1 + \frac{1}{T_1} \int_0^T v(t)\, dt\,.$$

Moreover, since the positions of the points are statistically

independent, the average of the product $\Pi[1 + v(t_j)]$ is just the product of the averages $\langle 1 + v(t_j)\rangle$, so that

$$L_T[v(t)] = \left[1 + \frac{1}{T_1}\int_0^T v(t)\,dt\right]^N.$$

Example 2. If in Example 1, we increase the interval length T_1 and the number of points N in proportion, so that their ratio $N/T_1 = \beta$, which equals the average density of points, remains constant, then in the limit we have

$$L_T[v(t)] = \lim_{N\to\infty}\left[1 + \frac{\beta}{N}\int_0^T v(t)\,dt\right]^N = \exp\left\{\beta\int_0^T v(t)\,dt\right\}.$$

This generating functional corresponds to a "Poisson system of points", which will be studied in Sec. 3. Examples of more complicated generating functionals will be given in Sec. 4.

In the case of a system of random points with continuous statistical properties, the generating functional is related to the distribution functions by the formula

$$L_T[v(t)] = 1 + \sum_{s=1}^{\infty}\frac{1}{s!}\int_0^T \cdots \int_0^T f_s(t_1, \ldots, t_s)\, v(t_1) \ldots v(t_s)\, dt_1 \ldots dt_s. \tag{6.4}$$

For other systems of random points, the distribution functions may have delta-function singularities at points where arguments coincide; then, formula (6.4) can be used to define the distribution functions in terms of the generating functional (6.3).

The generating functional completely characterizes a system of random points, and in this respect, it plays a role similar to that of the characteristic functional (1.60) in the theory of random functions. In particular, if $[t', t'']$ is a subinterval of the interval $[0, T]$, then the generating functional $L_{t',t''}[v]$ can be obtained by setting $v(t)$ equal to zero outside the interval $[t', t'']$, just as in the case of the characteristic functional.

Together with the distribution functions f_s, it is convenient to introduce the *correlation functions* (of the system of random points)

$$g_1(t_1), \qquad g_2(t_1, t_2), \qquad g_3(t_1, t_2, t_3), \ldots, \tag{6.5}$$

defined by the formula

$$1 + \sum_{s=1}^{\infty} \frac{1}{s!} \sum_{\alpha,\ldots,\omega=1}^{r} f_s(t_\alpha, \ldots, t_\omega)\, z_\alpha \ldots z_\omega = \exp\left\{ \sum_{s=1}^{\infty} \frac{1}{s!} \sum_{\alpha,\ldots,\omega=1}^{r} g_s(t_\alpha, \ldots, t_\omega)\, z_\alpha \ldots z_\omega \right\}, \tag{6.6}$$

which holds for arbitrary $z_1, z_2, \ldots$ and arbitrary $t_1, t_2, \ldots$ in the domain of definition of the system of random points. The correlation functions (6.5) bear the same relation to the distribution functions (6.1) as the correlation functions (1.64) bear to the moment functions (1.63).[2] Thus, applying the formulas (1.69) to f_s and g_s, we obtain

$$\begin{aligned} f_1(t_1) &= g_1(t_1), \\ f_2(t_1, t_2) &= g_2(t_1, t_2) + g_1(t_1)\, g_1(t_2), \\ f_3(t_1, t_2, t_3) &= g_3(t_1, t_2, t_3) + 3\{g_1(t_1)\, g_2(t_2, t_3)\}_s + g_1(t_1)\, g_1(t_2)\, g_1(t_3), \\ &\ldots \qquad\qquad \ldots\ldots\ldots\ldots, \end{aligned} \tag{6.7}$$

where the symbol $\{\ldots\}_s$ has the meaning explained on p. 19.

A system of random points is said to be *stationary* if the distribution functions describing it depend only on the time differences:

$$f_s(t_1, t_2, \ldots, t_s) = f'_s(t_2 - t_1, \ldots, t_s - t_1).$$

The most commonly encountered case is that of *ergodic systems* of random points. These are stationary systems characterized by the fact that the correlations between the events consisting of points falling at various different time instants vanish as the time instants

[2] Cf. formulas (1.66) and (1.67).

are moved apart. In other words, the correlation functions of an ergodic system go to zero, i.e.,

$$g_s(t_1, \ldots, t_s) \to 0\,, \tag{6.8}$$

provided that at least one difference $t_i - t_j$ of the arguments is increased without limit ($|\, t_i - t_j \,| \to \infty$). This does not apply to the distribution functions, which under the same conditions reduce to products of lower-order distribution functions (just as in the case of ordinary moment functions, in the absence of statistical dependence).

Choosing the points $t_1, t_2, \ldots$ in (6.6) closer and closer together, and suitably adjusting the quantities $z_1, z_2, \ldots$, we can transform (6.4) into

$$L_T[v(t)] = \exp\left\{\sum_{s=1}^{\infty} \frac{1}{s!} \int_0^T \ldots \int_0^T g_s(t_1, \ldots, t_s)\, v(t_1) \ldots v(t_s)\, dt_1 \ldots dt_s\right\} \tag{6.9}$$

[cf. the derivation of the characteristic functional (1.68)]. From this it is clear that the correlation functions can be defined as functional derivatives of the logarithm of the generating functional, i.e.,

$$g_s(t_1, \ldots, t_s) = \frac{\delta^s \ln L_T[v(t)]}{\delta v(t_1) \ldots \delta v(t_s)}\bigg|_{v(t)\equiv 0}, \tag{6.10}$$

while on the other hand,

$$f_s(t_1, \ldots, t_s) = \frac{\delta^s L_T[v(t)]}{\delta v(t_1) \ldots \delta v(t_s)}\bigg|_{v(t)\equiv 0}. \tag{6.11}$$

Certain quantities and functions characterizing systems of random points can be expressed simply in terms of the generating function. For example, let n be the (random) number of points falling in the interval $[0, T]$. Then, the characteristic function of n is

$$\langle e^{inu} \rangle = L_T[e^{iu} - 1]\,. \tag{6.12}$$

The probability that no points at all fall in $[0, T]$ is just

$$\mathbf{P}\{n = 0\} = L_T[-1]\,. \tag{6.13}$$

Next, we present without proof[3] some formulas involving the *conditional distribution functions* $p_1(t_1; T)$, $p_2(t_1, t_2; T)$, ... Here, the function $p_s(t_1, ..., t_s; T)$ gives the probability

$$dP = [p_s(t_1, ..., t_s; T) + O(\Delta)]\, \Delta_1 \dots \Delta_s \qquad (6.14)$$

that just one random point falls in each of the intervals (6.2) and that $n = s$, i.e., precisely s points fall in the whole interval $[0, T]$. Unlike the "unconditional" distribution functions, the conditional distribution functions are no longer local statistical characteristics, since their values at the times $t_1, ..., t_s$ depend on the interval $[0, T]$ in question. In fact, the conditional distribution functions are related to the unconditional distribution functions by the formulas

$$p_s(t_1, ..., t_s; T) = \sum_{r=0}^{\infty} \frac{(-1)^r}{r!} \int_0^T \dots \int_0^T f_{s+r}(t_1, ..., t_{s+r})\, dt_{s+1} \dots dt_{s+r}. \qquad (6.15)$$

Dividing (6.15) by $s!$, the number of permutations of the points falling in $[0, T]$ (the points can be distributed in various ways among the elementary intervals $[t_1, t_1 + \Delta_1], ..., [t_s, t_s + \Delta t_s]$), and then integrating with respect to the variables $t_1, ..., t_s$, we obtain the probability that precisely s points fall in $[0, T]$:

$$\mathbf{P}\{n = s\} \equiv P_s = \frac{1}{s!} \int_0^T \dots \int_0^T p_s(t_1, ..., t_s; T)\, dt_1 \dots dt_s. \qquad (6.16)$$

The meaning of the *normalized* conditional probability density

$$w_s(t_1, ..., t_s) = \frac{p_s(t_1, ..., t_s; T)}{s! P_s} \qquad (s = 1, 2, ...) \qquad (6.17)$$

is quite close to that of an ordinary probability density.

Solving the inverse problem, we can find the distribution functions $f_1, f_2, ...$ from a knowledge of the distribution functions

[3] See P. I. Kuznetsov and R. L. Stratonovich, *On the mathematical theory of correlated random points*, Izv. Akad. Nauk SSSR, Ser. Mat., **20**, 167 (1956).

(6.17) and the probabilities $P_1, P_2, \ldots$ that various numbers of points fall in $[0, T]$. The resulting formulas

$$f_s(t_1, \ldots, t_s) = \sum_{r=0}^{\infty} \frac{(s+r)!}{r!} P_{s+r} \int_0^T \cdots \int_0^T w_{s+r}(t_1, \ldots, t_{s+r})\, dt_{s+1} \ldots dt_{s+r} \tag{6.18}$$

are the inverses of (6.15). If the number of points in n is not random, then (6.18) goes over into the expression

$$f_s(t_1, \ldots, t_s) = \frac{n!}{(n-s)!} \int_0^T \cdots \int_0^T w_n(t_1, \ldots, t_n)\, dt_{s+1} \ldots dt_n \quad \text{for} \quad s \leqslant n, \tag{6.19}$$

$$f_s(t_1, \ldots, t_s) = 0 \quad \text{for} \quad s > n.$$

2. Random Functions Constructed from Systems of Random Points

Using a given system of random points, we can construct random functions in various ways. For example, in many cases we are concerned with a random function defined as a sum of the form

$$\eta(z) = \sum_j Q(z, t_j), \tag{6.20}$$

where the t_j are random points, $Q(z, t')$ is some known function, and z denotes some set of arguments (which in particular might be the time t). Among all functions of the form (6.20), a special role is played by the random function

$$\xi(t) = \sum_j \delta(t - t_j), \tag{6.21}$$

which will be called the "random density function." The more general random function (6.20) can be expressed as a linear transformation of the random density function (6.21):

$$\eta(z) = \int Q(z, t')\, \xi(t')\, dt'. \tag{6.22}$$

Therefore, to find the statistical characteristics of the random

function (6.20), we should first find the corresponding characteristics of the random density function.

Just like any random function, (6.21) is completely described by its characteristic functional (1.60). If we take account of the properties of the delta function, we can write the characteristic functional as

$$\Theta_T[u(t)] = \left\langle \exp\left\{ i \int_0^T u(t)\, \xi(t)\, dt \right\} \right\rangle = \left\langle \exp\left\{ i \sum_{j=1}^{n} u(t_j) \right\} \right\rangle$$

or

$$\Theta_T[u(t)] = \left\langle \prod_{j=1}^{n} e^{iu(t_j)} \right\rangle. \tag{6.23}$$

Here, as in (6.3), n is the number of random points falling in the interval $[0, T]$, and $t_1, \ldots, t_n$ are the coordinates of the points. Comparing (6.23) with the formula (6.3) which defines the generating functional, we find the formula relating the two functionals:

$$\Theta_T[u(t)] = L_T[e^{iu(t)} - 1]. \tag{6.24}$$

Thus, if we know the generating functional, we can determine the characteristic functional, and hence the other characteristics of $\xi(t)$. In particular, we can determine the correlation functions of $\xi(t)$, which can be expressed in terms of the correlation functions (6.5) of the system of random points. The corresponding formulas can be obtained either by functional differentiation with respect to $u(t)$, followed by setting $u(t) \equiv 0$, or by using the expansions (1.68) and (6.9). For example, substituting (1.68) and (6.9) into (6.24), taking the logarithm, and retaining only linear and quadratic terms in the $u(t_i)$, we find that

$$i \int_0^T k_1(t_1)\, u(t_1)\, dt_1 + \frac{i^2}{2} \int_0^T \int_0^T k_2(t_1, t_2)\, u(t_1)\, u(t_2)\, dt_1\, dt_2 + \ldots$$

$$= \int_0^T g_1(t_1) [e^{iu(t_1)} - 1]\, dt_1$$

$$+ \frac{1}{2} \int_0^T \int_0^T g_2(t_1, t_2) [e^{iu(t_1)} - 1] [e^{iu(t_2)} - 1]\, dt_1\, dt_2 + \ldots$$

or

$$i\int_0^T k_1(t_1)\,u(t_1)\,dt_1 + \frac{i^2}{2}\int_0^T\int_0^T k_2(t_1, t_2)\,u(t_1)\,u(t_2) + \ldots$$
$$= i\int_0^T g_1(t_1)\,u(t_1)\,dt_1 + \frac{i^2}{2}\int_0^T g_1(t_1)\,u^2(t_1)\,dt_1 \qquad (6.25)$$
$$+ \frac{i^2}{2}\int_0^T\int_0^T g_2(t_1, t_2)\,u(t_1)\,u(t_2)\,dt_1\,dt_2 + \ldots$$

Then, equating the linear and quadratic terms separately, we obtain

$$\begin{aligned} k_1(t_1) &= g_1(t_1)\,, \\ k_2(t_1, t_2) &= g_1(t_1)\,\delta(t_1 - t_2) + g_2(t_1, t_2)\,. \end{aligned} \qquad (6.26)$$

Similarly, writing more terms in (6.25), we can derive formulas for the higher-order correlation functions, e.g.,

$$\begin{aligned} k_3(t_1, t_2, t_3) &= \delta(t_1 - t_2)\,\delta(t_1 - t_3)\,g_1(t_1) \\ &\quad + 3\{\delta(t_1 - t_2)\,g_2(t_1, t_3)\}_s + g_3(t_1, t_2, t_3)\,, \\ k_4(t_1, t_2, t_3, t_4) &= \delta(t_1 - t_2)\,\delta(t_1 - t_3)\,\delta(t_1 - t_4)\,g_1(t_1) \\ &\quad + 4\{\delta(t_1 - t_2)\,\delta(t_1 - t_3)\,g_2(t_1, t_4)\}_s \\ &\quad + 3\{\delta(t_1 - t_2)\,\delta(t_3 - t_4)\,g_1(t_1, t_3)\}_s \\ &\quad + 6\{\delta(t_1 - t_2)\,g_3(t_1, t_3, t_4)\}_s + g_4(t_1, t_2, t_3, t_4)\,, \\ &\quad \ldots\ldots\ldots\ldots\ldots\ldots \end{aligned} \qquad (6.27)$$

Here, as on p. 19, the symbol $\{\ldots\}_s$ denotes the symmetrizing operation.

It is not hard to see that substitution of the expansions (1.66) and (6.4) into (6.24) gives the same formulas for the relations between the moment functions and the distribution functions:

$$\begin{aligned} m_1(t_1) &= f_1(t_1)\,, \\ m_2(t_1, t_2) &= \delta(t_1 - t_2)\,f_1(t_1) + f_2(t_1, t_2)\,, \\ m_3(t_1, t_2, t_3) &= \delta(t_1 - t_2)\,\delta(t_1 - t_3)\,f_1(t_1) \\ &\quad + 3\{\delta(t_1 - t_2)\,f_2(t_1, t_3)\}_s + f_3(t_1, t_2, t_3)\,, \\ &\quad \ldots\ldots\ldots\ldots\ldots\ldots \end{aligned} \qquad (6.28)$$

From a knowledge of the correlation (or moment) functions of the random density function, we can find the correlation (or moment) functions of any other random process of the form (6.20), by repeatedly applying the linear transformation (6.22). It will be observed that we are dealing with a completely symmetric situation: On the one hand, the relation between the g_s and the f_s is the same as that between the k_s and the m_s; on the other hand, the relation between the g_s and the k_s is the same as that between the f_s and the m_s. Moreover, there is a deep similarity between the mathematical apparatus of the theory of random functions and that of the theory of random points; in particular, this leads to the appearance of symmetric sums of the same kind in both theories.

As already remarked, the results given here can be generalized to the case of a system of random points in a multidimensional space, by simply replacing t by the coordinates $z_1, \ldots, z_p$ of a random point in the space. For example, the multidimensional generalization of (6.26) is

$$\begin{aligned} k_1(z_1, \ldots, z_p) &= g_1(z_1, \ldots, z_p), \\ k_2(z_1, \ldots, z_p; z_1', \ldots, z_p') &= g_1(z_1, \ldots, z_p)\,\delta(z_1 - z_1') \ldots \delta(z_p - z_p') \\ &\quad + g_2(z_1, \ldots, z_p; z_1', \ldots, z_p'). \end{aligned} \tag{6.29}$$

3. Poisson Systems of Random Points. Shot Noise

A system of random points is said to be a *Poisson system* if there are no correlations (of any order) between the points, i.e., if the occurrence of one or several points (in certain places) does not change the probabilities of occurrence of the other points. In this case, all the correlation functions of the system of points vanish except $g_1(t_1)$, i.e.,

$$g_2(t_1, t_2) = 0, \qquad g_3(t_1, t_2, t_3) = 0, \ldots \tag{6.30}$$

According to (6.9), the generating functional then has the form

$$L_T[v(t)] = \exp\left\{\int_0^T g_1(t)\, v(t)\, dt\right\}. \tag{6.31}$$

If we expand the exponential in series and compare it with the expansion (6.4), we easily see that the distribution functions factor into products of one-dimensional functions

$$f_s(t_1, \ldots, t_s) = g_1(t_1) \ldots g_1(t_s). \tag{6.32}$$

Using (6.15), (6.16) and (6.32), we find that

$$P_s = \frac{1}{s!}\left[\int_0^T g_1(t)\,dt\right]^s \exp\left\{-\int_0^T g_1(t)\,dt\right\} \qquad (s = 0, 1, 2, \ldots). \tag{6.33}$$

The distribution (6.33) is called a *Poisson distribution*, and is often used to define Poisson systems of points. In the stationary case, the average density of the points is constant, and (6.33) becomes

$$P_s = \frac{(g_1 T)^s}{s!} e^{-g_1 T}. \tag{6.34}$$

If we have a large number of independent events which take place (in time) with a finite average density, then the occurrence times of the events have a Poisson distribution. For example, consider the equilibrium current flow in a diode. When a single electron is emitted by the cathode and reaches the anode, it creates an anode current pulse $i(t - t_j)$, where the position of the pulse on the time axis is determined by t_j, the random time at which the electron is emitted. Moreover

$$\int i(t - t_j)\,dt = e, \tag{6.35}$$

where e is the charge of the electron. Dealing similarly with the other electrons, we obtain a set of random emission times

$$\ldots, t_{j-1}, t_j, t_{j+1}, \ldots$$

which correspond to a system of random points lying on the time axis. We shall assume that different electrons are emitted by the cathode independently of each other, so that the emission times have a Poisson distribution. In fact, there could only be dependence between electrons emitted almost simultaneously by neighboring

parts of the cathode, but this dependence can be neglected, if we assume that the dimensions of the cathode appreciably exceed the possible "correlation distance."

In terms of the average anode current $\langle I_a \rangle$, the average number of electrons which are emitted per unit time is

$$g_1 = \frac{\langle I_a \rangle}{e}, \tag{6.36}$$

and this parameter completely determines the system of random points. The anode current

$$I_a(t) = \sum_j i(t - t_j) \tag{6.37}$$

is a superposition of separate pulses, and hence is a random function of the form (6.20), constructed from a system of random points. We now use the formulas of the preceding section to find the statistical characteristics of $I_a(t)$.

First we assume that the finite size of the transit time from cathode to anode can be neglected. If we assume that this time is infinitely small, then each separate pulse can be regarded as a delta function, i.e.,

$$i(t - t_j) = e\delta(t - t_j),$$

and then the current (6.37) has the form

$$I_a(t) = e \sum_j \delta(t - t_j) = e\xi(t), \tag{6.38}$$

where $\xi(t)$ is given by (6.21). To find the correlation function of the current, we set $g_2 = 0$ in formula (6.26), obtaining

$$k_\xi(\tau) = k_2(t, t + \tau) = g_1\delta(\tau), \tag{6.39}$$

which implies

$$\mathbf{K}[I_a, I_{a\tau}] = e\langle I_a \rangle \delta(\tau), \tag{6.40}$$

according to (6.38) and (6.36).

The Fourier transform of (6.40) gives the spectral density of the "fluctuating component" $I_a - \langle I_a \rangle$ of the anode current:

$$S[I_a - \langle I_a \rangle; \omega] = e^2 S[\xi - \langle \xi \rangle; \omega] = 2e\langle I_a \rangle . \tag{6.41}$$

According to (6.41), which is known as *Schottky's formula*, the spectral density of the anode current fluctuations is proportional to the charge of the electron. This is natural, since the fluctuations are caused by the fact that electrical charge is discrete (hence the term "shot noise").

Taking account of (6.27) and the fact that the correlations (6.30) vanish, we easily see that the higher-order correlation functions also have "delta-function shape":

$$\begin{aligned} \mathbf{K}[I_a(t_1), \ldots, I_a(t_s)] &= e^s g_1 \delta(t_1 - t_2) \ldots \delta(t_1 - t_s) \\ &= e^{s-1}\langle I_a \rangle \delta(t_1 - t_2) \ldots \delta(t_1 - t_s) . \end{aligned} \tag{6.42}$$

However, at high frequencies, corresponding to oscillations whose periods are comparable to the transit time, we have to take account of the duration and shape of the individual pulses, by replacing the delta function by some other function $G(t - t_j)$, which is determined by the detailed motion of the electron. If the jth pulse is

$$i(t - t_j) = eG(t - t_j) , \tag{6.43}$$

then the sum of all the pulses is

$$I_a(t) = e \int G(t - t') \, \xi(t') \, dt''. \tag{6.44}$$

This, together with formula (6.39) for the correlation function, implies that

$$\begin{aligned} \mathbf{K}[I_a, I_{a\tau}] &= e^2 \int G(t - s) \, G(t + \tau - s) \, g_1 \, ds \\ &= e\langle I_a \rangle \int G(\tau + \sigma) \, G(\sigma) \, d\sigma . \end{aligned} \tag{6.45}$$

Using (6.44), we can easily find the relation between the spectral

density of the current fluctuations and the spectral density of the random density function; the formula relating them is just

$$S[I_a - \langle I_a \rangle; \omega] = e^2 \, |F(i\omega)|^2 \, S[\xi - \langle \xi \rangle; \omega] , \tag{6.46}$$

where

$$F(i\omega) = \int e^{-i\omega\tau} G(\tau) \, d\tau . \tag{6.47}$$

Moreover, it follows from (6.39) that

$$S[\xi - \langle \xi \rangle; \omega] = 2g_1 , \tag{6.48}$$

and hence

$$S[I_a - \langle I_a \rangle; \omega] = 2e\langle I_a \rangle \, |F(i\omega)|^2 . \tag{6.49}$$

This formula differs from (6.41) in that it contains the extra factor $|F(i\omega)|^2$, which might be called the "frequency depression factor," for a reason that will soon be apparent. It is easily seen from (6.35) and (6.43) that

$$F(0) = \int G(\tau) \, d\tau = 1 . \tag{6.50}$$

Therefore, the formulas (6.41) and (6.48) coincide at zero frequency. At other frequencies, $|F(i\omega)| < 1$ and (6.49) gives a smaller value than (6.41) for the spectral density of the shot noise. If the electrons undergo uniform acceleration in the cathode-anode space, which is possible if the electrodes are planar and the space-charge density is small, then a simple calculation[4] shows that

$$\begin{aligned} F(i\omega) &= \frac{2}{(i\omega T)^2} [1 - (1 + i\omega T) \, e^{-i\omega T}] , \\ |F(i\omega)|^2 &= \frac{8}{(\omega T)^4} [1 + \tfrac{1}{2}(\omega T)^2 - \cos \omega T - \omega T \sin \omega T] , \end{aligned} \tag{6.51}$$

where T is the transit time.

[4] We use the fact that if the initial velocity of the electron can be neglected, then $T = d\sqrt{2m/eV}$ and $i(t) = (e/d)^2 (V/m) \, t = 2et/T^2$, where d is the spacing and V the voltage difference between the electrodes, and m is the mass of the electron.

★ So far, we have assumed that all the electrons give rise to pulses of the same form. However, taking into account the effect of the longitudinal component v_j of the initial velocity of the jth electron, we should write the jth pulse as

$$i(t - t_j) = eG(t - t_j, v_j), \tag{6.52}$$

where the pulse shape now depends on v_j. Adding up the pulses and introducing the two-dimensional random density function

$$\xi(t, v) = \sum_j \delta(t - t_j)\, \delta(v - v_j), \tag{6.53}$$

we find that

$$I_a(t) = e \int\int G(t - t', v)\, \xi(t', v)\, dt'\, dv. \tag{6.54}$$

Here, $G(\tau, v)$ is a completely known function of two variables, determined by the detailed motion of the electrons. Thus, we are now dealing with a random system of points in the (t, v) plane. Assuming as before that the "initial data" of the different electrons are independent, we can again regard the points as forming a Poisson system. According to (6.29), in the two-dimensional case we have

$$\mathbf{K}[\xi(t, v), \xi(t + \tau, v')] = g_1(t, v)\, \delta(\tau)\, \delta(v - v') \tag{6.55}$$

instead of (6.39), and

$$\langle I_a \rangle = e \int\int G(t - t', v)\, g_1(t', v)\, dt'\, dv = e \int g_1(v)\, dv, \tag{6.56}$$

since the density $g_1(t, v) = g_1(v)$ is independent of the time and

$$\int G(\tau, v)\, d\tau = 1$$

for any v. Letting $w(v)$ be the normalized probability density of the initial velocity, we can write $g_1(v)$ in the form

$$g_1(v) = \frac{\langle I_a \rangle}{e}\, w(v). \tag{6.57}$$

Then it follows from (6.54), (6.55) and (6.57) that

$$\begin{aligned} \mathbf{K}[I_a, I_{a\tau}] &= e^2 \int\int G(t - t', v)\, G(t + \tau - t', v)\, g_1(v)\, dt'\, dv \\ &= e\langle I_a \rangle \int\int G(\sigma, v)\, G(\tau + \sigma, v)\, w(v)\, d\sigma\, dv\,. \star \end{aligned} \tag{6.58}$$

4. Systems of Random Points Defined by the First Two Moments

In the theory of random functions, one sometimes considers only second-order correlations, i.e., only the first two moments of the random function. Moreover, the formulas (6.26) and their inverses

$$\begin{aligned} g_1(t) &= \langle \xi \rangle\,, \\ g_2(t_1, t_2) &= \langle \xi(t_1)\, \xi(t_2) \rangle - \langle \xi(t_1) \rangle \langle \xi(t_2) \rangle - \langle \xi(t_1) \rangle\, \delta(t_1 - t_2) \end{aligned} \tag{6.59}$$

show that the first two moments of a random function like (6.21) or (6.22), which is constructed from random points, are uniquely related to the first two distribution functions of the system of points. Therefore, in a theory involving only the first two moments, we can regard the functions g_1, g_2 or f_1, f_2 as known. In general, however, the first two distribution functions do not completely characterize the system of random points, and hence, to describe a system of random points more completely, it may be necessary to consider higher-order moments. Nevertheless, in some special cases this is not necessary, and the system of random points, and also random functions constructed from it, are completely determined by the first two moments. We now consider two such cases.

4.1. Systems of points correlated in pairs. Given a system of random points, we say that the points are *correlated in pairs* if all the correlation functions except g_1 and g_2 vanish, i.e., if

$$g_3 = g_4 = \ldots = 0\,.$$

Then the generating functional (6.9) has the form

$$L_T[v(t)] = \exp\left\{ \int_0^T g_1(t)\, v(t)\, dt + \frac{1}{2} \int_0^T \int_0^T g_2(t, t')\, v(t)\, v(t')\, dt\, dt' \right\}. \tag{6.60}$$

4.2. Systems of nonapproaching points. Suppose that all the correlation functions can be expressed in terms of two functions $f_1(t)$ and $R(t, t')$ by the formulas

$$g_n(t_1, \ldots, t_n) = (-1)^{n-1}(n-1)! f_1(t_1) \ldots f_1(t_n) \{R(t_2, t_1) \ldots R(t_n, t_1)\}_s \tag{6.61}$$

where, as before, $\{\ldots\}_s$ denotes the symmetrizing operation. Then, according to (6.9), the generating functional associated with these correlation functions is

$$L_T[v(t)] = \exp\left\{\int_0^T \frac{\ln\left[1 + \int_0^T R(t, t') f_1(t') v(t')\, dt'\right]}{\int_0^T R(t, t') f_1(t') v(t')\, dt'} f_1(t) v(t)\, dt\right\}. \tag{6.62}$$

Here we have used the formula

$$\sum_{s=1}^{\infty} \frac{1}{s}(-1)^{s-1} I^{s-1} = I^{-1} \ln(1 + I),$$

in carrying out the summation.

The meaning of the function $R(t, t')$ is revealed by setting $s = 2$ in (6.61); this gives

$$g_2(t_1, t_2) = -f_1(t_1) f_1(t_2) R(t_1, t_2)$$

or

$$f_2(t_1, t_2) = f_1(t_1) f_1(t_2) [1 - R(t_1, t_2)]. \tag{6.63}$$

It is appropriate to call $R(t, t')$ the *correlation coefficient*, since (6.63) shows that the distribution function f_2 vanishes when $R = 1$. Moreover, when $R \approx 1$, (6.62) becomes

$$L_T[v(t)] \approx 1 + \int_0^T f_1(t) v(t)\, dt,$$

which, according to (6.4), means that all the distribution functions vanish, except the first. If τ_{cor} characterizes how fast the correlation falls off, so that $R(t, t') \approx 1$ for small time separations $t - t' \ll \tau_{cor}$, then it is clear that our random points cannot lie near each other, i.e., the intervals between random points cannot be much less than

τ_{cor}. Therefore, the random points with correlation functions (6.61) are called a *system of nonapproaching points*. In particular, using the expression (6.62), we can find the probability that no points at all will fall in the interval $[0, T]$. According to (6.13), we have

$$\mathbf{P}\{n = 0\} = \exp\left\{\int_0^T \frac{\ln\left[1 - \int_0^T R(t, t')\, f_1(t')\, dt'\right]}{\int_0^T R(t, t')\, f_1(t')\, dt'} f_1(t)\, dt\right\}. \qquad (6.64)$$

In many problems, in order to obtain concrete results, it is expedient to substitute systems of nonapproaching points, of the type just described, for the more complicated systems of random points actually encountered. For example, formulas from the theory of systems of nonapproaching points will be used in Vol. II, Chap. 3, Sec. 2, to calculate the distribution of the durations of up-crossings (peaks) of fluctuation noise.

5. Spectral Densities of Various Pulse Sequences

In this section, we shall calculate the spectral densities of various kinds of pulse sequences. We begin by considering a sequence of pulses of "standard form," described by the function $G(t)$. In general, the positions of the different pulses along the time axis are random, so that the pulse sequence has the form

$$\eta(t) = \sum_j G(t - t_j), \qquad (6.65)$$

where the t_j are random points. We can regard (6.65) as a linear transformation of the random density function (6.21), i.e.,

$$\eta(t) = \int_{-\infty}^{\infty} G(t - t')\, \xi(t')\, dt', \qquad (6.66)$$

or in spectral form,

$$\eta_\omega = F(i\omega)\, \xi_\omega, \qquad (6.67)$$

where

$$F(i\omega) = \int_{-\infty}^{\infty} e^{-i\omega\tau} G(\tau)\, d\tau\,,$$

$$\eta_\omega = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega t} \eta(t)\, dt\,,$$

$$\xi_\omega = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega t} \xi(t)\, dt\,.$$

We shall assume that the pulse sequence (6.65) is a stationary random process, which means that the system of random points giving the pulse positions is also stationary. Then it follows from (6.66) of (6.67) that

$$\begin{aligned} \langle \eta \rangle &= F(0)\, \langle \xi \rangle\,, \\ S[\eta;\, \omega] &= |\, F(i\omega)\,|^2\, S[\xi;\, \omega]\,. \end{aligned} \tag{6.68}$$

But according to (2.9), (2.12) and the second of the equations (6.26) or (6.28), the spectral density $S[\xi;\, \omega]$ can easily be expressed in terms of the functions g_1 and g_2, i.e.,

$$\begin{aligned} S[\xi;\, \omega] &= 2g_1 + 2\gamma(\omega) + 4\pi g_1^2 \delta(\omega)\,, \\ S[\xi - \langle \xi \rangle;\, \omega] &= 2[\gamma(\omega) + g_1]\,, \end{aligned} \tag{6.69}$$

where

$$\gamma(\omega) = \int_{-\infty}^{\infty} e^{i\omega\tau} g_2(\tau)\, d\tau\,. \tag{6.70}$$

Substituting (6.69) into (6.68), we obtain

$$\begin{aligned} S[\eta;\, \omega] &= 4\pi F^2(0)\, g_1^2 \delta(\omega) + 2\,|\, F(i\omega)\,|^2\, [\gamma(\omega) + g_1]\,, \\ S[\eta - \langle \eta \rangle;\, \omega] &= 2\,|\, F(i\omega)\,|^2\, [\gamma(\omega) + g_1]\,. \end{aligned} \tag{6.71}$$

In the special case where the function $G(\tau)$ is a rectangular pulse of fixed height G_0 and width τ_p (see Figure 1), so that

$$G(\tau) = \begin{cases} G_0 & \text{for } 0 < \tau < \tau_0\,, \\ 0 & \text{for } \tau < 0 \text{ or } \tau > \tau_0\,, \end{cases} \tag{6.72}$$

we have

$$F(i\omega) = \frac{G_0}{i\omega}[1 - e^{-i\omega\tau_0}], \tag{6.73}$$

and then (6.68) and (6.71) become

$$\begin{aligned} \langle\eta\rangle &= G_0\tau_0 g_1, \\ S[\eta - \langle\eta\rangle; \omega] &= \frac{8G_0^2}{\omega^2}\sin^2\frac{\omega\tau_0}{2}[\gamma(\omega) + g_1]. \end{aligned} \tag{6.74}$$

Thus, we see that there is no particular difficulty in finding the spectral density of a pulse sequence when we know the first two functions g_1 and g_2. However, if g_1 and g_2 are not known, we have to calculate the spectral density of the pulse sequence by a different method, and then use (6.69) and (6.70) to find g_1, γ and g_2 instead. We now consider an example of this type.

Let the random points t_j be such that the distances between neighboring points are independent identically distributed random variables, with a given probability distribution $w(\zeta)$. We choose $t = 0$ as the time origin, and we renumber the random points with positive coordinates $t_j > 0$ in such a way that t_0 has the smallest coordinate. Thus, we have

$$t_1 - t_0 = \zeta_1, \qquad t_2 - t_1 = \zeta_2, \ldots.$$

or

$$t_1 = t_0 + \zeta_1, \quad t_2 = t_0 + \zeta_1 + \zeta_2, \ldots, \quad t_j = t_0 + \zeta_1 + \ldots + \zeta_j, \ldots. \tag{6.75}$$

as in Figure 1.

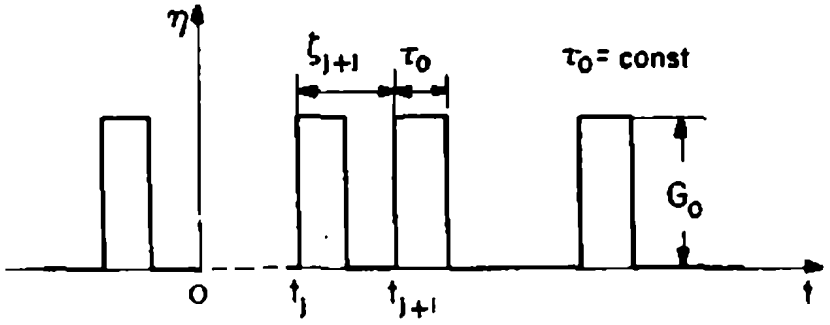

Fig. 1. A pulse sequence constructed from the random points (6.75), where each pulse has height G_0 and width τ_0.

To calculate the spectral density of the random function (6.21), we use formula (2.27), which requires that we pass to the limit $\epsilon \to 0$ in the expression

$$\epsilon \left\langle \left| L\left[\xi; \frac{\epsilon}{2} - i\omega\right] \right|^2 \right\rangle,$$

where

$$L[\xi; p] = \int_0^\infty e^{-pt}\, \xi(t)\, dt = e^{-pt_0}\left[1 + e^{-p\zeta_1} + e^{-p\zeta_1 - p\zeta_2} + \ldots\right] \tag{6.76}$$

is the Laplace transform of the random function $\xi(t)$. It follows that

$$|\, L[\xi; p]\,|^2 = e^{-(p+p^*)t_0} \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} e^{-p(\zeta_1+\ldots+\zeta_j)-p^*(\zeta_1+\ldots+\zeta_k)}, \tag{6.77}$$

where the asterisk denotes the complex conjugate. Next, we divide the terms in this double sum into two groups: The first group contains the terms for which $j \geqslant k$, and we write $j - k = l$; the second group contains the remaining terms, for which $j < k$, and we write $k - j = l$. Then (6.77) becomes

$$\begin{aligned} \langle |\, L[\xi; p]\,|^2 \rangle = \mathrm{e}^{-(p+p^*)t_0} \Bigg\{ & \sum_{k=1}^{\infty} \sum_{l=0}^{\infty} e^{-(p+p^*)(\zeta_1+\ldots+\zeta_k)-p(\zeta_{k+1}+\ldots+\zeta_{k+l})} \\ & + \sum_{j=1}^{\infty} \sum_{l=1}^{\infty} e^{-(p+p^*)(\zeta_1+\ldots+\zeta_j)-p^*(\zeta_{j+1}+\ldots+\zeta_{j+l})} \Bigg\}. \end{aligned} \tag{6.78}$$

To calculate the mean value of (6.78), we use the independence of the different random variables ζ_j, from which it follows that every term of the double sum can be written as a product of characteristic functions corresponding to the distribution $w(\zeta)$, taken with different arguments. Thus, writing

$$\Theta(u) = \int_0^\infty e^{iu\zeta}\, w(\zeta)\, d\zeta = \langle e^{iu\zeta} \rangle \tag{6.79}$$

(recall that ζ can take only positive values), we have

$$\begin{aligned} \langle e^{-p\zeta} \rangle &= \Theta(ip), \\ \langle e^{-p^*\zeta} \rangle &= \Theta(ip^*) = \Theta^*(ip). \end{aligned} \tag{6.80}$$

Therefore, the mean value of (6.78) takes the form

$$\langle |L[\xi; p]|^2 \rangle = \langle e^{-(p+p^*)t_0} \rangle \left\{ \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \Theta^k(ip + ip^*)\, \Theta^l(ip) + \sum_{j=1}^{\infty} \sum_{l=1}^{\infty} \Theta^j(ip + ip^*)\, \Theta^{*l}(ip) \right\}. \tag{6.81}$$

Summing the series in (6.81) [which is easily done, since they are geometric series], multiplying by 2ϵ, and then setting $p = \frac{1}{2}\epsilon - i\omega$, we obtain

$$\left\langle \left| L\left[\xi; \frac{\epsilon}{2} - i\omega\right] \right|^2 \right\rangle = \frac{\langle e^{-st_0} \rangle\, \Theta(i\epsilon)}{1 - \Theta(i\epsilon)} \left\{ \frac{1}{1 - \Theta[\omega + (i\epsilon/2)]} + \frac{\Theta^*[\omega + (i\epsilon/2)]}{1 - \Theta^*[\omega + (i\epsilon/2)]} \right\}. \tag{6.82}$$

Finally, according to (2.27), to find the spectral density we need only pass to the limit $\epsilon \to 0$. If $\omega \neq 0$, $\Theta(\omega) \neq 1$, we immediately find that

$$S[\xi; \omega] = \frac{2}{\langle \zeta \rangle} \left\{ \frac{1}{1 - \Theta(\omega)} + \frac{\Theta^*(\omega)}{1 - \Theta^*(\omega)} \right\} = \frac{2}{\langle \zeta \rangle} \frac{1 - |\Theta(\omega)|^2}{|1 - \Theta(\omega)|^2}, \tag{6.83}$$

since by (1.22),

$$\frac{1 - \Theta(i\epsilon)}{\epsilon} \to \langle \zeta \rangle.$$

For the frequency range lying near the origin $\omega = 0$, the situation is more complicated, and a separate argument is needed. In this case, we use the first few terms of the expansion

$$\Theta\left(\omega + \frac{i\epsilon}{2}\right) = 1 + i\left(\omega + \frac{i\epsilon}{2}\right)\langle \zeta \rangle + O\left[\left(\omega + \frac{i\epsilon}{2}\right)^2\right], \tag{6.84}$$

which gives

$$\begin{aligned} &\frac{1}{1 - \Theta[\omega + (i\epsilon/2)]} + \frac{\Theta^*[\omega + (i\epsilon/2)]}{1 - \Theta^*[\omega + (i\epsilon/2)]} \\ &\quad = \frac{1}{\langle \zeta \rangle} \left[\frac{1}{(\epsilon/2) - i\omega} + \frac{1}{(\epsilon/2) + i\omega} \right] \left[1 + O\left(\omega + \frac{i\epsilon}{2}\right) \right] \\ &\quad = \frac{1}{\langle \zeta \rangle} \frac{\epsilon}{(\epsilon/2)^2 + \omega^2} \left[1 + O\left(\omega + \frac{i\epsilon}{2}\right) \right]. \end{aligned} \tag{6.85}$$

We now use the relation

$$\frac{\epsilon}{(\epsilon/2)^2 + \omega^2} \to 2\pi\delta(\omega) \quad \text{as} \quad \epsilon \to 0\,, \tag{6.86}$$

which is easily verified directly by integrating both sides. Then, taking account of (6.82), (6.85) and (6.86), we see that the spectral density $S[\xi; \omega]$ has a delta-function singularity at the origin. Adding this singularity to formula (6.83), which is valid for $\omega \neq 0$, we obtain

$$S[\xi; \omega] = \frac{2}{\langle\zeta\rangle} \frac{1 - |\,\Theta(\omega)\,|^2}{|\,1 - \Theta(\omega)\,|^2} + \frac{4\pi\delta(\omega)}{\langle\zeta\rangle^2}\,. \tag{6.87}$$

The presence of the delta-function singularity

$$\frac{4\pi\delta(\omega)}{\langle\zeta\rangle^2}$$

in (6.87) is due to the fact that the mean value $\langle\xi\rangle$ is nonzero. In fact, $\langle\xi\rangle$ is just the reciprocal of the mean distance $\langle\zeta\rangle$ between the random points, i.e.,

$$\langle\xi\rangle = \langle\zeta\rangle^{-1}\,. \tag{6.88}$$

If we center the random function $\xi(t)$ about its mean value, then, according to (2.12) and (6.87),

$$S[\xi - \langle\xi\rangle; \omega] = \frac{2}{\langle\zeta\rangle} \frac{1 - |\,\Theta(\omega)\,|^2}{|\,1 - \Theta(\omega)\,|^2}\,. \tag{6.89}$$

It is easily verified that (6.89) can also be written in the form

$$S[\xi - \langle\xi\rangle; \omega] = \frac{2}{\langle\zeta\rangle} \operatorname{Re} \frac{1 + \Theta(\omega)}{1 - \Theta(\omega)}\,. \tag{6.90}$$

From a knowledge of the spectral density of the random function $\xi(t)$, we can immediately calculate the spectral density of the pulse sequence (6.65). Thus, from (6.90) and the second of the formulas (6.68), we obtain

$$S[\eta - \langle\eta\rangle; \omega] = \frac{2}{\langle\zeta\rangle} |\,F(i\omega)\,|^2 \operatorname{Re} \frac{1 + \Theta(\omega)}{1 - \Theta(\omega)}\,. \tag{6.91}$$

In the special case of identical strictly rectangular pulses of height G_0 and width τ_0, it follows from (6.73) that

$$S[\eta - \langle\eta\rangle; \omega] = \frac{8G_0^2}{\omega^2\langle\zeta\rangle} \sin^2 \frac{\omega\tau_0}{2} \frac{1 - |\Theta(\omega)|^2}{|1 - \Theta(\omega)|^2}. \tag{6.92}$$

These formulas allow us to determine the Fourier transform $\gamma(\omega)$ of the correlation function $g_2(\tau)$ of the system of random points under consideration. Comparing (6.69) and (6.90) and using (6.88), we find that

$$\gamma(\omega) = \frac{2}{\langle\zeta\rangle} \operatorname{Re} \frac{\Theta(\omega)}{1 - \Theta(\omega)}. \tag{6.93}$$

Then, taking the inverse Fourier transform of (6.70), we find that

$$g_2(\tau) = \frac{1}{\pi\langle\zeta\rangle} \int_{-\infty}^{\infty} e^{i\omega\tau} \operatorname{Re} \frac{\Theta(\omega)}{1 - \Theta(\omega)}\, d\omega = \frac{1}{\pi\langle\zeta\rangle} \int_{-\infty}^{\infty} \frac{\Theta(\omega)}{1 - \Theta(\omega)} \cos\omega\tau\, d\omega. \tag{6.94}$$

★ Using the formulas just derived, we can find the connection between the characteristic function $\Theta(\omega)$ and the Laplace transform $L[k(\tau); p]$ of the correlation function $k_\xi(\tau) \equiv k(\tau)$. Differentiating both sides of (2.14) with respect to τ, we find that

$$\frac{dk(\tau)}{d\tau} = -\frac{1}{4\pi} \int_{-\infty}^{\infty} e^{-i\omega\tau} i\omega S[\xi - \langle\xi\rangle; \omega]\, d\omega. \tag{6.95}$$

Then, we substitute (6.90) into (6.95), obtaining

$$\begin{aligned} \frac{dk(\tau)}{d\tau} = &-\frac{1}{4\pi\langle\zeta\rangle} \int_{-\infty}^{\infty} e^{-i\omega\tau} i\omega \frac{1 + \Theta(\omega)}{1 - \Theta(\omega)}\, d\omega \\ &-\frac{1}{4\pi\langle\zeta\rangle} \int_{-\infty}^{\infty} e^{-i\omega\tau} i\omega \frac{1 + \Theta(-\omega)}{1 - \Theta(-\omega)}\, d\omega. \end{aligned} \tag{6.96}$$

The second term in the right-hand side can be written in the form

$$\frac{1}{4\pi\langle\zeta\rangle} \int_{-\infty}^{\infty} e^{i\omega'\tau} i\omega' \frac{1 + \Theta(\omega')}{1 - \Theta(\omega')}\, d\omega' \qquad (\omega' = -\omega),$$

which is just the negative of the first term in the right-hand side, except that τ has been replaced by $-\tau$. It follows from (6.79) that

$$\Theta(ip) = \int_0^\infty e^{-p\zeta} w(\zeta)\, d\zeta\,,$$

i.e., $\Theta(ip)$ is the Laplace transform of the function $w(\zeta)$. Moreover, $\Theta(ip)$ is analytic for $\operatorname{Re} p \geqslant 0$, and the integral

$$w(\zeta) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega\zeta}\Theta(\omega)\, d\omega = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} e^{p\zeta}\Theta(ip)\, dp \qquad (p = -i\omega)$$

is just the inverse Laplace transform.

Now consider the function

$$p\,\frac{1 + \Theta(ip)}{1 - \Theta(ip)}\,, \tag{6.97}$$

which is also analytic for $\operatorname{Re} p \geqslant 0$, provided that ζ is not a deterministic quantity.[5] Clearly, (6.97) is the Laplace transform of the function

$$\frac{1}{2\pi i}\int_{-i\infty}^{i\infty} e^{p\tau} p\,\frac{1 + \Theta(ip)}{1 - \Theta(ip)}\, dp = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega\tau}(-i\omega)\,\frac{1 + \Theta(\omega)}{1 - \Theta(\omega)}\, d\omega\,,$$

which vanishes for $\tau < 0$. Therefore, the first term in the right-hand side of (6.96) vanishes for $\tau < 0$ and the second term vanishes if $\tau > 0$, so that (6.96) can be written in the form

$$L\left[\frac{dk(\tau)}{d\tau}\,;\, p\right] = \frac{p}{2\langle\zeta\rangle}\,\frac{1 + \Theta(ip)}{1 - \Theta(ip)}$$

for $\tau > 0$. It follows from

$$k(\tau) = -\int_\tau^\infty \frac{dk}{d\tau}\, d\tau = \int_0^\tau \frac{dk}{d\tau}\, d\tau - \int_0^\infty \frac{dk}{d\tau}\, d\tau$$

that

$$L[k(\tau); p] = \frac{1}{p}\left\{L\left[\frac{dk}{d\tau}\,;\, p\right] - L\left[\frac{dk}{d\tau}\,;\, 0\right]\right\}, \tag{6.98}$$

[5] If $\zeta = c$ with probability 1, where c is a constant, then (6.97) will have poles on the imaginary axis, i.e., for $\operatorname{Re} p = 0$.

and hence (6.97) implies

$$L[k(\tau); p] = \frac{1}{2\langle\zeta\rangle} \frac{1 + \Theta(ip)}{1 - \Theta(ip)} - \frac{1}{p\langle\zeta\rangle^2}. \tag{6.99}$$

A similar relation can also be written for the function (6.94):

$$L[g_2(\tau); p] = \frac{1}{\langle\zeta\rangle} \frac{\Theta(ip)}{1 - \Theta(ip)} - \frac{1}{p\langle\zeta\rangle^2}. \star \tag{6.100}$$

Next, we consider a sequence of pulses which are not identical. Specifically, let the sequence consist of strictly rectangular pulses, all of unit height, whose lengths vary from pulse to pulse, so that both the initial points and the end points of the pulses form systems of random points (see Figure 2). Thus, the random points on the time axis are now of two kinds. Actually, the problem could be reduced to the study of a system of random points of one kind, which lie in a plane rather than on a line. However, we shall not give the general theory here, but instead confine ourselves to the special problem just described.

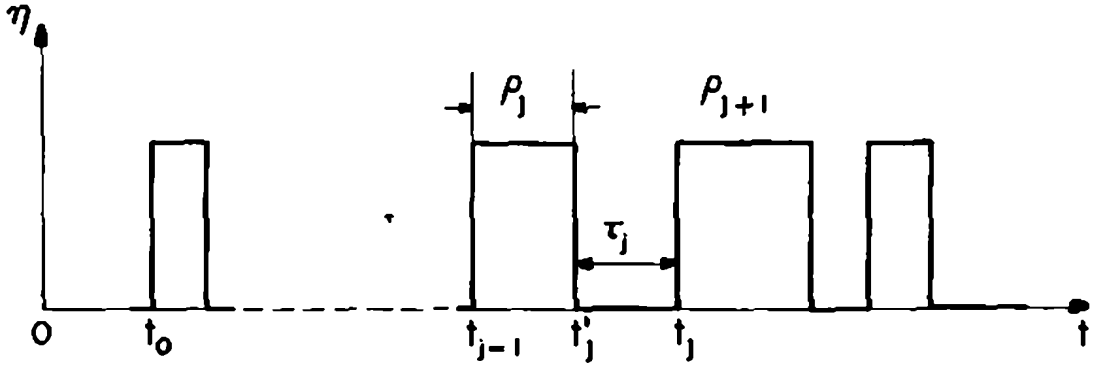

Fig. 2. A sequence of pulses with random positions and random lengths.

Thus, let the initial points t_j and the points t'_j of the pulses be numbered in order of increasing time:

$$\ldots < t_0 < t'_1 < t_1 < t'_2 < \ldots \tag{6.101}$$

Then, the pulse sequence can be written as

$$\eta(t) = \sum_j \vartheta_{t'_{j+1}-t_j}(t - t_j), \tag{6.102}$$

where

$$\vartheta_\tau(t) = \begin{cases} 1 & \text{for } 0 < t < \tau, \\ 0 & \text{for } t < 0 \text{ and } t > 0. \end{cases}$$

Instead of the function (6.102), it is more convenient to consider its derivative

$$\dot{\eta}(t) = \xi(t) = \sum_j [\delta(t - t_j) - \delta(t - t_j')], \tag{6.103}$$

which generalizes the random density function (6.21) to the case of nonidentical points. If we wish to find the spectral density of the stationary pulse sequence (6.102), we have to use the fact that the spectral densities of $\xi(t)$ and $\eta(t)$ are related by the formula

$$S[\eta; \omega] = \frac{1}{\omega^2} S[\xi; \omega] \tag{6.104}$$

or

$$S[\eta - \langle\eta\rangle; \omega] = \frac{1}{\omega^2} S[\xi - \langle\xi\rangle; \omega]. \tag{6.105}$$

We shall only consider the case where the following two conditions are met:

1. The pulse lengths

$$t_j' - t_{j-1} = \rho_j \tag{6.106}$$

are independent, identically distributed random variables, with characteristic function

$$\Theta_1(u) = \langle e^{iu\rho_j} \rangle \tag{6.107}$$

for any j;

2. The intervals between pulses

$$t_j - t_j' = \sigma_j \tag{6.108}$$

are independent, identically distributed random variables, with characteristic function

$$\Theta_2(u) = \langle e^{iu\sigma_j} \rangle \tag{6.109}$$

for any j.

Just as before, we use formula (2.27) to calculate the spectral density $S[\xi; \omega]$. Of all the initial times t_j, let t_0 have the smallest positive value. Then the Laplace transform of the function (6.103) can be written in the form

$$L[\xi; p] = e^{-pt_0} - e^{-pt_1'} + e^{-pt_1} - e^{-pt_2'} - \ldots \tag{6.110}$$

This is the appropriate formula for the case where the origin $t = 0$ lies in an interval between two pulses. In the opposite case, where the origin $t = 0$ is covered by a pulse, we have to write

$$L[\xi; p] = -e^{-pt_0'} + e^{-pt_0} - e^{-pt_1'} + e^{-pt_1} - \ldots \tag{6.111}$$

There are definite probabilities P_1 and P_2 for each of the two relations (6.110) and (6.111) to hold; for the moment, we shall only consider (6.110). Taking (6.106) and (6.108) into account, we can write (6.110) in the form

$$\begin{aligned} L[\xi; p] &= e^{-pt_0}[1 - e^{-p\rho_1} + e^{-p\rho_1 - p\sigma_1} - e^{-p\rho_1 - p\sigma_1 - p\rho_2} + \ldots] \\ &= e^{-pt_0} \sum_{j=0}^{\infty} \{\exp[-p(\rho_1 + \sigma_1 + \ldots + \rho_j + \sigma_j)] \\ &\qquad - \exp[-p(\rho_1 + \sigma_1 + \ldots + \sigma_j + \rho_{j+1})]\}. \end{aligned} \tag{6.112}$$

To calculate $\langle |L[\xi; p]^r|^2 \rangle$, we multiply (6.110) or (6.112) by its complex conjugate. This introduces the double sum

$$\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} A_j A_k^*,$$

where

$$A_j = e^{-pt_j} - e^{-pt_{j+1}'},$$

which we then divide into three separate sums, corresponding to $j = k$, $j > k$ and $j < k$. Thus, we write

$$|L[\xi; p]|^2 = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} A_j A_k^* = \sum_0 + \sum_+ + \sum_-, \tag{6.113}$$

where

$$\sum_0 = \sum_j A_j A_j^*, \quad \sum_+ = \sum_{j>k} A_j A_k^* = \sum_{k=0}^{\infty} \sum_{l=1}^{\infty} A_{k+l} A_k^*, \quad \sum_- = \sum_+^*.$$

In more detailed form, we have

$$\begin{aligned} \sum_0 = e^{-(p+p^*)t_0} \sum_{j=0}^{\infty} \{ & \exp[-p(\rho_1 + \sigma_1 + \ldots + \rho_j + \sigma_j)] \\ & - \exp[-p(\rho_1 + \sigma_1 + \ldots + \sigma_j + \rho_{j+1})]\} \\ & \times \{\exp[-p^*(\rho_1 + \sigma_1 + \ldots + \rho_j + \sigma_j)] \\ & - \exp[-p^*(\rho_1 + \sigma_1 + \ldots + \sigma_j + \rho_{j+1})]\}, \end{aligned} \tag{6.114}$$

and also

$$\begin{aligned} \sum_+ = {} & e^{-(p+p^*)t_0} \\ & \times \sum_{k=0}^{\infty} \sum_{l=1}^{\infty} \{\exp[-p(\rho_1 + \ldots + \rho_{k+l} + \sigma_{k+l}) - p^*(\rho_1 + \ldots + \rho_k + \sigma_k)] \\ & - \exp[-p(\rho_1 + \ldots + \rho_{k+l} + \sigma_{k+l}) - p^*(\rho_1 + \ldots + \sigma_k + \rho_{k+1})] \\ & - \exp[-p(\rho_1 + \ldots + \sigma_{k+l} + \rho_{k+l+1}) - p^*(\rho_1 + \ldots + \rho_k + \sigma_k)] \\ & + \exp[-p(\rho_1 + \ldots + \sigma_{k+l} + \rho_{k+l+1}) - p^*(\rho_1 + \ldots + \sigma_k + \rho_{k+1})]\}. \end{aligned} \tag{6.115}$$

Averaging the last two expressions, we obtain products of the characteristic functions (6.107) and (6.109), since the random variables ρ_j and σ_j are independent. Thus, (6.114) leads to the expression

$$\begin{aligned} \left\langle \sum_0 \right\rangle = \langle e^{-(p+p^*)t_0} \rangle \sum_{j=0}^{\infty} & \Theta_1^j(ip + ip^*)\, \Theta_2^j(ip + ip^*) \\ & \times [1 - \Theta_1(ip) - \Theta_1(ip^*) + \Theta_1(ip + ip^*)]. \end{aligned}$$

Summing the geometric series in the right-hand side and setting $p + p^* = \epsilon$, we obtain

$$\left\langle \sum_0 \right\rangle = \frac{\langle e^{-\epsilon t_0} \rangle}{1 - \Theta_1(i\epsilon)\, \Theta_2(i\epsilon)} [1 + \Theta_1(i\epsilon) - \Theta_1(ip) - \Theta_1^*(ip)]. \tag{6.116}$$

Similarly, averaging (6.115) leads to the expression

$$\left\langle \sum_{+} \right\rangle = \langle e^{-\epsilon t_0} \rangle \sum_{k=0}^{\infty} \sum_{l=1}^{\infty} [\Theta_1^k(i\epsilon)\, \Theta_2^k(i\epsilon)\, \Theta_1^l(ip)\, \Theta_2^l(ip)$$
$$- \Theta_1^{k+1}(i\epsilon)\, \Theta_2^k(i\epsilon)\, \Theta_1^{l-1}(ip)\, \Theta_2^l(ip) - \Theta_1^k(i\epsilon)\, \Theta_2^k(i\epsilon)\, \Theta_1^{l+1}(ip)\, \Theta_2^l(ip)$$
$$+ \Theta_1^{k+1}(i\epsilon)\, \Theta_2^k(i\epsilon)\, \Theta_1^l(ip)\, \Theta_2^l(ip)] .$$

Factoring out $\Theta_1^k(i\epsilon)\Theta_2^k(i\epsilon)\Theta_1^l(ip)\Theta_2^l(ip)$, and summing the geometric series in k and l, we obtain

$$\left\langle \sum_{+} \right\rangle = \frac{\langle e^{-\epsilon t_0} \rangle}{1 - \Theta_1(i\epsilon)\, \Theta_2(i\epsilon)} \frac{\Theta_1(ip)\, \Theta_2(ip)}{1 - \Theta_1(ip)\, \Theta_2(ip)}$$
$$\times \left[1 - \frac{\Theta_1(i\epsilon)}{\Theta_1(ip)} - \Theta_1(ip) + \Theta_1(i\epsilon)\right] . \tag{6.117}$$

According to (2.27) and (6.113), the spectral density we are looking for can be found from the formula

$$\frac{1}{2} S[\xi; \omega] = \lim_{\epsilon \to 0} \epsilon \left\langle \sum_{0} \right\rangle + \lim_{\epsilon \to 0} \epsilon \left\langle \sum_{+} \right\rangle + \lim_{\epsilon \to 0} \epsilon \left\langle \sum_{-} \right\rangle$$
$$= \lim_{\epsilon \to 0} \epsilon \left\langle \sum_{0} \right\rangle + 2 \operatorname{Re} \lim_{\epsilon \to 0} \epsilon \left\langle \sum_{+} \right\rangle , \tag{6.118}$$

after setting

$$p = \frac{\epsilon}{2} - i\omega .$$

If $\omega \neq 0$, $\Theta_1(\omega)\Theta_2(\omega) \neq 1$, the passage to the limit in (6.118) is trivial. In fact, since

$$\Theta_1(i\epsilon) = 1 - \epsilon\langle \rho \rangle + \tfrac{1}{2} \epsilon^2 \langle \rho^2 \rangle + \dots ,$$
$$\Theta_2(i\epsilon) = 1 - \epsilon\langle \sigma \rangle + \tfrac{1}{2} \epsilon^2 \langle \sigma^2 \rangle + \dots ,$$

we have

$$\frac{\epsilon}{1 - \Theta_1(i\epsilon)\, \Theta_2(i\epsilon)} \to \frac{1}{\langle \rho \rangle + \langle \sigma \rangle} \text{ as } \epsilon \to 0 . \tag{6.119}$$

Therefore, it follows from (6.116) and (6.117) that

$$\lim_{\epsilon\to 0}\epsilon\left\langle \sum_0 \right\rangle = \frac{1}{\langle\rho\rangle + \langle\sigma\rangle}\left[2 - \Theta_1(\omega) - \Theta_1^*(\omega)\right]$$

$$= \frac{2}{\langle\rho\rangle + \langle\sigma\rangle}\operatorname{Re}\left[1 - \Theta_1(\omega)\right], \tag{6.120}$$

$$\lim_{\epsilon\to 0}\epsilon\left\langle \sum_+ \right\rangle = \frac{1}{\langle\rho\rangle + \langle\sigma\rangle}\frac{\Theta_1(\omega)\Theta_2(\omega)}{1 - \Theta_1(\omega)\Theta_2(\omega)}\left[2 - \frac{1}{\Theta_1(\omega)} - \Theta_1(\omega)\right].$$

A special argument shows that the relations (6.120) remain the same in the range $\omega \approx 0$, $\Theta_1(\omega)\Theta_2(\omega) \approx 1$. Substituting (6.120) into (6.118), we obtain the formula

$$\frac{1}{2}S[\xi;\omega] = \frac{2}{\langle\rho\rangle + \langle\sigma\rangle}\operatorname{Re}\frac{\left[1 - \Theta_1(\omega)\right]\left[1 - \Theta_2(\omega)\right]}{1 - \Theta_1(\omega)\,\Theta_2(\omega)}. \tag{6.121}$$

In the above derivation, we used the expressions which are valid for the case where the time origin $t = 0$ lies in an interval between two pulses. In the opposite case, where the time origin is covered by a pulse, we have to use (6.111) instead of (6.110), making all appropriate subsequent changes. Thus, to obtain the final result, we have to average two different expressions of the form (6.118), where the first comes from (6.110) and occurs with probability P_1, while the second comes from (6.111) and occurs with probability P_2. However, it is easily seen that replacing (6.110) by (6.111) does not change the expressions obtained when we pass to the limit $\epsilon \to 0$. Therefore, the need for the additional averaging disappears, and (6.121) is actually the final result.

To find the spectral density of the pulse sequence (6.102), we use formula (6.104), obtaining

$$\frac{1}{2}S[\eta;\omega] = \frac{1}{\omega^2}\frac{2}{\langle\rho\rangle + \langle\sigma\rangle}\operatorname{Re}\frac{\left[1 - \Theta_1(\omega)\right]\left[1 - \Theta_2(\omega)\right]}{1 - \Theta_1(\omega)\,\Theta_2(\omega)}. \tag{6.122}$$

There remains a certain arbitrariness in the definition of the value of $S[\eta;\omega]$ for $\omega \approx 0$. An analysis of the expression (6.121) reveals the following behavior for small frequencies

$$\operatorname{Re}\frac{\left[1 - \Theta_1(\omega)\right]\left[1 - \Theta_2(\omega)\right]}{1 - \Theta_1(\omega)\,\Theta_2(\omega)} = \frac{\omega^2}{2}\frac{\langle\rho\rangle^2\mathbf{D}\sigma + \langle\sigma\rangle^2\mathbf{D}\rho}{\left[\langle\rho\rangle + \langle\sigma\rangle\right]^2}. \tag{6.123}$$

Therefore, (6.122) tends to a finite limit as $\omega \to 0$. If we take this limit as the definition of the function $\omega^{-2}S[\xi; \omega]$ for $\omega = 0$, we obtain the spectral density of the process $\eta(t) - \langle\eta\rangle$:

$$S[\eta - \langle\eta\rangle; \omega] = 4\omega^{-2}[\langle\rho\rangle + \langle\sigma\rangle]^{-1} \operatorname{Re} \frac{[1 - \Theta_1(\omega)]\,[1 - \Theta_2(\omega)]}{1 - \Theta_1(\omega)\,\Theta_2(\omega)},$$
$$S[\eta - \langle\eta\rangle; 0] = 2\,\frac{\langle\rho\rangle^2 \mathbf{D}\sigma + \langle\sigma\rangle^2 \mathbf{D}\rho}{[\langle\rho\rangle + \langle\sigma\rangle]^3}. \tag{6.124}$$

We can derive a variety of special results from (6.124). For example, suppose the pulse length is a fixed number τ_0, so that

$$\Theta_2(\omega) = e^{i\omega\tau_0}, \qquad \Theta_1(\omega)\,\Theta_2(\omega) = \Theta(\omega) = \langle e^{i\omega\zeta}\rangle, \tag{6.125}$$

where

$$\zeta = \rho + \sigma.$$

Then, from (6.124) we obtain

$$\begin{aligned} S[\eta - \langle\eta\rangle; \omega] &= \frac{4}{\omega^2\langle\zeta\rangle} \operatorname{Re}\left[(1 - \Theta)^{-1}(1 - e^{-i\omega\tau_0}\Theta)\,(-2i)\,e^{i\omega\tau_0/2} \sin\frac{\omega\tau_0}{2}\right] \\ &= \frac{8}{\omega^2\langle\zeta\rangle} \sin^2\frac{\omega\tau_0}{2}\,\frac{1 - |\,\Theta\,|^2}{|\,1 - \Theta\,|^2}, \end{aligned} \tag{6.126}$$

which agrees with (6.92) for $G_0 = 1$. Moreover, if both the pulse lengths and the lengths of the intervals between pulses have the same distribution, i.e., if

$$\Theta_1(\omega) = \Theta_2(\omega) = \Theta(\omega), \tag{6.127}$$

then formula (6.124) gives

$$S[\eta - \langle\eta\rangle; \omega] = \frac{2}{\omega^2\langle\rho\rangle} \operatorname{Re}\frac{1 - \Theta(\omega)}{1 + \Theta(\omega)} = \frac{2}{\omega^2\langle\rho\rangle}\,\frac{1 - |\,\Theta(\omega)\,|^2}{|\,1 + \Theta(\omega)\,|^2}. \tag{6.128}$$

Sometimes it is convenient to use formula (6.124) after first writing it in the form

$$S[\eta - \langle\eta\rangle; \omega] = \frac{4}{\omega^2}\,\frac{1}{\langle\rho\rangle + \langle\sigma\rangle} \operatorname{Re}\left[\frac{1}{1 - \Theta_1(\omega)} + \frac{1}{1 - \Theta_2(\omega)} - 1\right]^{-1}. \tag{6.129}$$

★ By the same method as used to derive formula (6.99), we can use formula (6.124) to derive the Laplace transform of the correlation function $k_\eta(\tau)$. Writing (6.124) as a sum of two complex conjugate expressions, substituting into (6.95), and taking account of the fact that one of the resulting integrals vanishes, we find that

$$\frac{dk_\eta(\tau)}{d\tau} = \frac{1}{2\pi\langle\rho + \sigma\rangle}\int_{-\infty}^{\infty} e^{-i\omega\tau}\,\frac{1}{i\omega}\,\frac{[1 - \Theta_1(\omega)]\,[1 - \Theta_2(\omega)]}{1 - \Theta_1(\omega)\,\Theta_2(\omega)}\,d\omega$$

for $\tau > 0$. It follows that

$$L\left[\frac{dk_\eta(\tau)}{d\tau}; p\right] = -\frac{1}{\langle\rho + \sigma\rangle}\,\frac{1}{p}\,\frac{[1 - \Theta_1(ip)]\,[1 - \Theta_2(ip)]}{1 - \Theta_1(ip)\,\Theta_2(ip)}. \tag{6.130}$$

Then, using (6.98), we obtain

$$L[k_\eta(\tau); p] = -\frac{1}{\langle\rho + \sigma\rangle}\,\frac{1}{p^2}\,\frac{[1 - \Theta_1(ip)]\,[1 - \Theta_2(ip)]}{1 - \Theta_1(ip)\,\Theta_2(ip)} + \frac{\langle\rho\rangle\,\langle\sigma\rangle}{\langle\rho + \sigma\rangle^2}\,\frac{1}{p}. \star \tag{6.131}$$

CHAPTER 7

Narrow-Band Random Processes

1. Equivalence of a Narrow-Band Process to a Pair of Slowly Varying Processes

By a *narrow-band process*, we mean a stationary random process whose realizations are close to being sinusoidal oscillations of some fixed frequency ω_0, for time intervals equal to a large number of periods $2\pi/\omega_0$; clearly, such a process has mean value zero. Figure 3 shows the typical appearance of a narrow-band process.

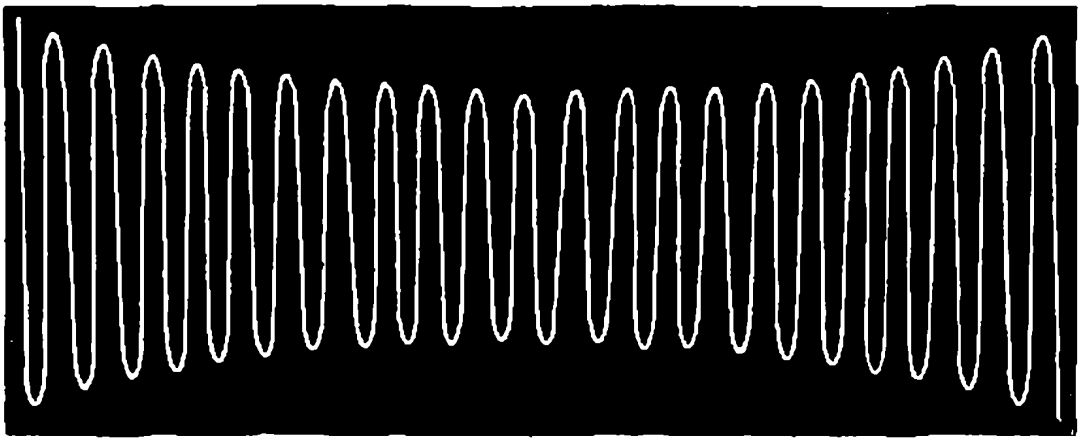

Fig. 3. A narrow-band process.

In spectral language, a narrow-band process is described by the fact that its spectral density is negligibly small everywhere except in a narrow frequency band

$$\omega_0 - \frac{\Delta\omega}{2} \leqslant \omega \leqslant \omega_0 + \frac{\Delta\omega}{2},$$

where $\Delta\omega \ll \omega_0$. If $\Delta\omega$ is the smallest such "bandwidth," then the time constant

$$T_1 = \frac{1}{\Delta\omega},$$

which is much larger than the oscillation period $2\pi/\omega_0$, characterizes the time it takes the amplitude and phase of the process $y(t)$ to change appreciably, i.e., deviations between the narrow-band process $y(t)$ and the ideal sinusoid it approximates begin to "accumulate" during a time of order T_1.

Suppose we are given two stationary random processes $y_1(t)$ and $y_2(t)$, with zero mean values, whose rates of change are characterized by the time constant T_1, and suppose we form the new process

$$x(t) = y_1(t)\cos\omega_0 t - y_2(t)\sin\omega_0 t\,. \tag{7.1}$$

Then, the process $x(t)$ will be narrow-band if $\omega_0 \gg 1/T_1$, i.e., if the frequency ω_0 is so large that the functions $y_1(t)$ and $y_2(t)$ do not manage to change appreciably during the course of a period $2\pi/\omega_0$. To obtain the correlation function of $x(t)$, we multiply (7.1) by the similar expression corresponding to time $t+\tau$, and then average the result, obtaining

$$\begin{aligned}\langle xx_\tau\rangle = {} & \langle y_1y_{1\tau}\rangle\cos\omega_0 t\cos\omega_0(t+\tau) - \langle y_1y_{2\tau}\rangle\cos\omega_0 t\sin\omega_0(t+\tau)\\ & - \langle y_2y_{1\tau}\rangle\sin\omega_0 t\cos\omega_0(t+\tau) + \langle y_2y_{2\tau}\rangle\sin\omega_0 t\sin\omega_0(t+\tau)\end{aligned}$$

or

$$\begin{aligned}\langle xx_\tau\rangle = {} & \tfrac{1}{2}[\langle y_1y_{1\tau}\rangle\cos\omega_0\tau - \langle y_1y_{2\tau}\rangle\sin\omega_0\tau\\ & + \langle y_2y_{1\tau}\rangle\sin\omega_0\tau + \langle y_2y_{2\tau}\rangle\cos\omega_0\tau]\\ & + \tfrac{1}{2}[\langle y_1y_{1\tau}\rangle\cos\omega_0(2t+\tau) - \langle y_1y_{2\tau}\rangle\sin\omega_0(2t+\tau)\\ & - \langle y_2y_{1\tau}\rangle\sin\omega_0(2t+\tau) - \langle y_2y_{2\tau}\rangle\cos\omega_0(2t+\tau)]\,.\end{aligned} \tag{7.2}$$

It is clear from (7.2) that in general the process $x(t)$ is nonstationary, since $\langle xx_\tau\rangle$ depends not only on τ, but also on t. In fact, process $x(t)$ is stationary if and only if

$$\begin{aligned}\langle y_1y_{1\tau}\rangle &= \langle y_2y_{2\tau}\rangle\,,\\ \langle y_1y_{2\tau}\rangle &= -\langle y_2y_{1\tau}\rangle\,.\end{aligned}$$

Assuming that these relations hold, we introduce the notation

$$\begin{aligned} \langle y_1 y_{1\tau} \rangle &= \langle y_2 y_{2\tau} \rangle = r(\tau) , \\ \langle y_1 y_{2\tau} \rangle &= s(\tau) , \end{aligned} \tag{7.3}$$

where $s(-\tau) = -s(\tau)$. Then, (7.2) becomes

$$\langle x x_\tau \rangle = r(\tau) \cos \omega_0 \tau - s(\tau) \sin \omega_0 \tau . \tag{7.4}$$

Writing (7.4) in the form

$$\langle x x_\tau \rangle = \tfrac{1}{2} (r + is) e^{i\omega_0 \tau} + \tfrac{1}{2} (r - is) e^{-i\omega_0 \tau} ,$$

and taking the Fourier transform of this expression, we find that according to (2.9),

$$S[x; \omega] = \rho(\omega + \omega_0) - \sigma(\omega + \omega_0) + \rho(\omega - \omega_0) + \sigma(\omega - \omega_0) , \tag{7.5}$$

where

$$\begin{aligned} \rho(\Omega) &= \int_{-\infty}^{\infty} e^{i\Omega\tau} r(\tau)\, d\tau = 2 \int_0^{\infty} r(\tau) \cos \Omega\tau \, d\tau , \\ \sigma(\Omega) &= -i \int_{-\infty}^{\infty} e^{i\Omega\tau} s(\tau)\, d\tau = 2 \int_0^{\infty} s(\tau) \sin \Omega\tau \, d\tau . \end{aligned} \tag{7.6}$$

It is easy to see that the first of the functions (7.6) is even, while the second is odd:

$$\rho(-\Omega) = \rho(\Omega) , \qquad \sigma(-\Omega) = -\sigma(\Omega) . \tag{7.7}$$

Since the functions $\rho(\Omega)$ and $\sigma(\Omega)$ are appreciably different from zero only in the range $\Omega \sim \Delta\omega \cong 1/T_1$, and since $\omega_0 \gg \Delta\omega$, the functions $\rho(\omega + \omega_0)$ and $\sigma(\omega + \omega_0)$ are small for positive frequencies $\omega > 0$. Neglecting these terms in (7.5), we find that

$$S[x; \omega] \approx \rho(\omega - \omega_0) + \sigma(\omega - \omega_0) \qquad (\omega > 0) . \tag{7.8}$$

The relations (7.7) allow us to solve (7.8) for $\rho(\Omega)$ and $\sigma(\Omega)$:

$$\begin{aligned} \rho(\Omega) &\approx \tfrac{1}{2} S[x, \omega_0 + \Omega] + \tfrac{1}{2} S[x, \omega_0 - \Omega] , \\ \sigma(\Omega) &\approx \tfrac{1}{2} S[x, \omega_0 + \Omega] - \tfrac{1}{2} S[x, \omega_0 - \Omega] . \end{aligned} \tag{7.9}$$

Thus, the functions $\rho(\Omega)$ and $\sigma(\Omega)$ uniquely determine the spectral density $S[x;\omega]$, and similarly the correlation functions $r(\tau)$ and $s(\tau)$ uniquely determine the correlation function $\langle xx_\tau\rangle$, and conversely.

Instead of two real processes $y_1(t)$ and $y_2(t)$, we can consider one complex slowly varying process

$$z(t) = y_1(t) + iy_2(t), \tag{7.10}$$

which has the correlation function

$$\begin{aligned}\langle zz_\tau^*\rangle &= \langle y_1y_{1\tau}\rangle + \langle y_2y_{2\tau}\rangle - i\langle y_1y_{2\tau}\rangle + i\langle y_2y_{1\tau}\rangle \\ &= 2[r(\tau) - is(\tau)]\end{aligned} \tag{7.11}$$

and the spectral density

$$S[z;\Omega] = 4[\rho(\Omega) + \sigma(\Omega)]. \tag{7.12}$$

Comparing (7.12) and (7.8), we see that the spectral densities of the processes $x(t)$ and $z(t)$ are related by the formula

$$S[x;\omega] \approx \tfrac{1}{4} S[z;\omega - \omega_0]. \tag{7.13}$$

If we multiply $z(t)$ by the harmonic time factor $e^{i\omega_0 t}$, we obtain a rapidly varying process

$$\tilde{z}(t) = z(t)\, e^{i\omega_0 t}, \tag{7.14}$$

which has $x(t)$ as its real part:

$$x(t) = \operatorname{Re}\tilde{z}(t). \tag{7.15}$$

A comparison of (7.4) and (7.11) shows at once that

$$\langle xx_\tau\rangle = \tfrac{1}{2}\operatorname{Re}[\langle zz_\tau^*\rangle\, e^{-i\omega_0\tau}], \tag{7.16}$$

which is equivalent to the following formula relating the correlation functions of the real and complex narrow-band processes:

$$\langle xx_\tau\rangle = \tfrac{1}{2}\operatorname{Re}\langle z\tilde{z}_\tau^*\rangle.$$

One can also consider the imaginary part of $\tilde{z}(t)$, i.e., the function

$$\bar{x}(t) = \operatorname{Im} \tilde{z}(t) = y_1(t) \sin \omega_0 t + y_2(t) \cos \omega_0 t , \tag{7.17}$$

which is called the *conjugate process* of $x(t)$ [not to be confused with the complex conjugate]. It is easily verified that

$$\begin{aligned} \langle \bar{x}\bar{x}_\tau \rangle &= r(\tau) \cos \omega_0 \tau - s(\tau) \sin \omega_0 \tau , \\ \langle x\bar{x}_\tau \rangle &= r(\tau) \sin \omega_0 \tau + s(\tau) \cos \omega_0 \tau , \end{aligned} \tag{7.18}$$

while the spectral density of the conjugate process and the cross-spectral density of the original process and its conjugate are given by

$$\begin{aligned} &S[\bar{x}; \omega] = S[x; \omega] , \\ &S[x, \bar{x}; \omega] = -i\rho(\omega + \omega_0) + i\sigma(\omega + \omega_0) + i\rho(\omega - \omega_0) + i\sigma(\omega - \omega_0). \end{aligned} \tag{7.19}$$

Because of the narrow-band condition ($\omega_0 \gg \Delta\omega$), we can write the last formula as

$$S[x, \bar{x}; \omega] \approx \begin{cases} i\rho(\omega - \omega_0) + i\sigma(\omega - \omega_0) \approx iS[x; \omega] & \text{for } \omega > 0 , \\ -i\rho(-\omega - \omega_0) - i\sigma(-\omega - \omega_0) \approx -iS[x; -\omega] & \text{for } \omega < 0 . \end{cases} \tag{7.20}$$

The formulas (7.1) and (7.17) establish a one-to-one correspondence between the two rapidly varying processes $x(t)$, $\bar{x}(t)$ and the two slowly varying processes $y_1(t)$, $y_2(t)$. In fact, solving (7.1) and (7.17) for $y_1(t)$ and $y_2(t)$, we find that

$$\begin{aligned} y_1(t) &= x(t) \cos \omega_0 t + \bar{x}(t) \sin \omega_0 t , \\ y_2(t) &= -x(t) \sin \omega_0 t + \bar{x}(t) \cos \omega_0 t . \end{aligned} \tag{7.21}$$

As is easily seen from (7.3) and (7.18), the processes $x(t)$ and $\bar{x}(t)$ satisfy the same relations as the processes $y_1(t)$ and $y_2(t)$:

$$\begin{aligned} \langle xx_\tau \rangle &= \langle \bar{x}\bar{x}_\tau \rangle \\ \langle x\bar{x}_\tau \rangle &= -\langle x\bar{x}_{-\tau} \rangle = -\langle x_\tau \bar{x} \rangle . \end{aligned} \tag{7.22}$$

It can also be shown that the formulas (7.22) are necessary conditions for the slowly varying processes (7.21) to be stationary, provided that $x(t)$ and $\bar{x}(t)$ are stationary.

We have just studied the narrow-band process (7.1) constructed from two given slowly varying processes $y_1(t)$ and $y_2(t)$. Next, we consider the opposite problem of finding two slowly varying processes $y_1(t)$ and $y_2(t)$ corresponding to a given narrow-band process. According to (7.21), the definition of $y_1(t)$ and $y_2(t)$ depends on the definition of the conjugate process $\tilde{x}(t)$. Suppose $\tilde{x}(t)$ is given by a stationary (i.e., time-invariant) linear transformation

$$\tilde{x}(t) = \int_{-\infty}^{\infty} G(t - t')\, x(t')\, dt' , \tag{7.23}$$

applied to the original narrow-band process $x(t)$. In spectral form, we can write (7.23) as

$$\tilde{x}_\omega = F(i\omega) x_\omega , \tag{7.24}$$

where

$$x_\omega = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega t} x(t)\, dt ,$$

$$\tilde{x}_\omega = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega t} \tilde{x}(t)\, dt ,$$

$$F(i\omega) = \int_{-\infty}^{\infty} e^{-i\omega\tau} G(\tau)\, d\tau .$$

Then, obviously we have

$$\begin{aligned} S[\tilde{x}, \omega] &= |F(i\omega)|^2 S[x; \omega] , \\ S[x, \tilde{x}; \omega] &= F^*(i\omega) S[x; \omega] . \end{aligned} \tag{7.25}$$

Comparing these formulas with (7.19) and (7.20), we see that $F(i\omega)$ has to satisfy the relations

$$F(i\omega) = \begin{cases} -i & \text{for} \quad \omega - \omega_0 \sim \Delta\omega , \\ i & \text{for} \quad \omega + \omega_0 \sim \Delta\omega . \end{cases} \tag{7.26}$$

These relations guarantee that (7.20) and (7.22) hold, which in turn, together with (7.21), allow us to prove (7.4), (7.18) and the other results given above.

Formula (7.26) describes the behavior of the function $F(i\omega)$ only in the narrow intervals $|\omega \pm \omega_0| \sim \Delta\omega$ where the spectral density $S[x; \omega]$ is appreciably different from zero. Elsewhere, the choice of $F(i\omega)$ is arbitrary. The important point is that (7.23) must transform $\sin \omega_0 t$ into $-\cos \omega_0 t$ and $\cos \omega_0 t$ into $\sin \omega_0 t$. We now show two possible specific forms of the transformation (7.23) and of the function $F(i\omega)$ which have particular advantages.

If we choose

$$F(i\omega) = -i \operatorname{sgn} \omega , \tag{7.27}$$

where

$$\operatorname{sgn} \omega = \frac{\omega}{|\omega|} ,$$

then (7.23) becomes the *Hilbert transform*

$$\tilde{x}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(t')}{t - t'} dt' = \frac{1}{\pi} \int_0^{\infty} \frac{x(t - \tau - x(t + \tau)}{\tau} d\tau . \tag{7.28}$$

In this case, we can write the complex process (7.14) as

$$\begin{aligned} \check{z}(t) = x(t) + i\tilde{x}(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\omega t} x_\omega \, d\omega + i \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\omega t} F(i\omega) \, x_\omega \, d\omega \\ &= \sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{i\omega t} x_\omega \, d\omega . \end{aligned} \tag{7.29}$$

This way of defining $F(i\omega)$ has the advantage that it can be generalized immediately to the case of processes which are not narrow-band.

However, the integral operation (7.28) is not convenient unless we are dealing with problems in spectral form. Therefore, we now indicate another way of defining the conjugate process, which involves differentiation instead of an integral operation, i.e.,

$$\tilde{x}(t) = -\frac{1}{\omega_0} \dot{x}(t) , \qquad F(i\omega) = -\frac{i\omega}{\omega_0} . \tag{7.30}$$

This way of defining $\tilde{x}(t)$ is convenient for making the generalization to nonstationary processes and for studying situations lying outside

the context of correlation theory, i.e., situations where we use the apparatus of Markov processes or the theory of higher moments. (Only the first two moments are considered in the correlation theory.) The advantage of (7.30) is that it can be used in cases where $x(t)$ is described by a differential equation.

It follows from (7.30) that the complex processes (7.10) and (7.14) can be written as

$$\tilde{z}(t) = x + i\tilde{x} = x + \frac{\dot{x}}{i\omega_0}, \qquad z(t) = \left(x + \frac{\dot{x}}{i\omega_0}\right) e^{-i\omega_0 t}. \tag{7.31}$$

Differentiating the second of these equations with respect to t, we obtain

$$\dot{z} = \frac{1}{i\omega_0} e^{-i\omega_0 t}(\ddot{x} + \omega_0^2 x), \tag{7.32}$$

an expression which is particularly convenient when the process $x(t)$ is described by a differential equation of the form

$$\ddot{x} + \omega_0^2 x = \omega_0^2 f(x, \dot{x}, t).$$

Then (7.32) gives

$$\dot{z} = -i\omega_0 e^{-i\omega_0 t} f(x, \dot{x}, t), \tag{7.33}$$

where

$$\begin{aligned} x &= \frac{1}{2}(z e^{i\omega_0 t} + z^* e^{-i\omega_0 t}), \\ \dot{x} &= \frac{i\omega_0}{2}(z e^{i\omega_0 t} - z^* e^{-i\omega_0 t}). \end{aligned} \tag{7.34}$$

To summarize, in this section we have shown that a single narrow-band process $x(t)$ is equivalent to two slowly varying processes $y_1(t)$ and $y_2(t)$. Moreover, we have shown how to go from $x(t)$ to $y_1(t)$, $y_2(t)$, and *vice versa*, by using linear operations.

2. Narrow-Band Processes Defined by Differential Equations

In radio engineering, narrow-band processes are formed in frequency-selective circuits (containing capacitances, inductances and resistances), which are described by differential equations.

Moreover, the signal at the output of such a circuit will be narrow-band even when the external or internal fluctuations acting on the circuit are broad-band. Therefore, we shall consider differential equations containing a rapidly varying random process. We begin by studying the second-order linear equation

$$\ddot{x} + \epsilon\dot{x} + x = \epsilon\xi(t), \tag{7.35}$$

describing the behavior of a resonant circuit under the influence of the noise $\xi(t)$. Here, ϵ is a small parameter which guarantees that the narrow-band condition $\Delta\omega \ll 1$ is satisfied (note that $\omega_0 = 1$). In fact, it can easily be seen that the condition $\Delta\omega \ll 1$ is equivalent to the condition $\epsilon \ll 1$.

According to (7.21) and (7.30), the slowly varying random functions y_1 and y_2 are just

$$\begin{aligned} y_1 &= x\cos t - \dot{x}\sin t, \\ y_2 &= -x\sin t - \dot{x}\cos t. \end{aligned} \tag{7.36}$$

Differentiating y_1 and y_2 with respect to t, we obtain

$$\begin{aligned} \dot{y}_1 &= -(\ddot{x} + x)\sin t, \\ \dot{y}_2 &= -(\ddot{x} + x)\cos t, \end{aligned}$$

or

$$\begin{aligned} \dot{y}_1 &= \epsilon(\dot{x} - \xi)\sin t, \\ \dot{y}_2 &= \epsilon(\dot{x} - \xi)\cos t, \end{aligned}$$

where we have substituted from (7.35). Here, $\dot{x}$ should be expressed in terms of y_1 and y_2 by using the formulas

$$\begin{aligned} x &= y_1\cos t - y_2\sin t, \\ -\dot{x} &= y_1\sin t + y_2\cos t, \end{aligned}$$

which are the inverses of (7.36). As a result, instead of the single equation (7.35), we obtain the system of equations

$$\begin{aligned} \dot{y}_1 &= -\frac{\epsilon}{2}y_1 - \epsilon\xi\sin t + \frac{\epsilon}{2}y_1\cos 2t - \frac{\epsilon}{2}y_2\sin 2t, \\ \dot{y}_2 &= -\frac{\epsilon}{2}y_2 - \epsilon\xi\cos t - \frac{\epsilon}{2}y_1\sin 2t - \frac{\epsilon}{2}y_2\cos 2t. \end{aligned} \tag{7.37}$$

The appearance of the small parameter ϵ on the right in (7.37) confirms that the processes y_1 and y_2 change slowly. We shall say that the equations (7.37) are in *standard form*, and we shall refer to the double-frequency terms

$$\frac{\epsilon}{2} y_1 \cos 2t\,, \qquad \frac{\epsilon}{2} y_2 \sin 2t\,, \qquad \frac{\epsilon}{2} y_1 \sin 2t\,, \qquad \frac{\epsilon}{2} y_2 \cos 2t$$

in (7.37) as the *oscillatory terms*. The effect of these terms during a "period" of the processes y_1 and y_2 is small and can be neglected in the first approximation; then, the equations (7.37) reduce to

$$\begin{aligned} \dot{y}_1 &= -\frac{\epsilon}{2} y_1 - \epsilon\xi \sin t\,, \\ \dot{y}_2 &= -\frac{\epsilon}{2} y_2 - \epsilon\xi \cos t\,. \end{aligned} \tag{7.38}$$

If ϵ is sufficiently small, we can investigate the equations (7.38) by using the method given in Chap. 4, Secs. 7-9; this corresponds to replacing the actual processes $y_1(t)$ and $y_2(t)$ by Markov processes. Using this method, we can replace (7.38) by the equivalent "simplified" equations

$$\begin{aligned} \dot{y}_1 &= -\frac{\epsilon}{2} y_1 - \epsilon\xi_1\,, \\ \dot{y}_2 &= -\frac{\epsilon}{2} y_2 - \epsilon\xi_2\,, \end{aligned} \tag{7.39}$$

where ξ_1 and ξ_2 are delta-correlated random processes with zero mean values and correlation functions

$$\begin{aligned} \langle \xi_1 \xi_{1\tau} \rangle &= \delta(\tau) \int_{-\infty}^{\infty} \langle \xi \xi_\tau \rangle \sin t \sin (t + \tau)\, d\tau\,, \\ \langle \xi_2 \xi_{2\tau} \rangle &= \delta(\tau) \int_{-\infty}^{\infty} \langle \xi \xi_\tau \rangle \cos t \cos (t + \tau)\, d\tau\,, \\ \langle \xi_1 \xi_{2\tau} \rangle &= \delta(\tau) \int_{-\infty}^{\infty} \langle \xi \xi_\tau \rangle \sin t \cos (t + \tau)\, d\tau\,. \end{aligned} \tag{7.40}$$

Each of the integrals appearing in (7.40) can be written as a sum of two integrals, e.g.,

$$\int_{-\infty}^{\infty} \langle \xi \xi_\tau \rangle \sin t \sin (t + \tau)\, d\tau = \frac{1}{2}\int_{-\infty}^{\infty} \langle \xi \xi_\tau \rangle \cos \tau\, d\tau - \frac{1}{2}\int_{-\infty}^{\infty} \langle \xi \xi_\tau \rangle \cos (2t + \tau)\, d\tau .$$

Here, the second integral in the right-hand side can be neglected in the first approximation, since it represents a nonstationary oscillatory contribution, which has only a slight effect on the Fokker-Planck equation when averaged over a period. The same is true of the other two integrals in (7.40), from which the oscillatory parts should also be subtracted. As a result, instead of (7.40), we have

$$\begin{gathered} \langle \xi_1 \xi_{1\tau} \rangle = \langle \xi_2 \xi_{2\tau} \rangle = K_0 \delta(\tau) , \\ \langle \xi_1 \xi_{2\tau} \rangle = 0 , \end{gathered} \tag{7.41}$$

where

$$K_0 = \frac{1}{2}\int_{-\infty}^{\infty} \langle \xi \xi_\tau \rangle \cos \tau\, d\tau = \frac{1}{4} S[\xi; 1] . \tag{7.42}$$

It follows from (7.39) and (7.41) that $y_1(t)$ and $y_2(t)$ are identical exponentially correlated random processes. In Chap. 4, Sec. 10, it was shown that the correlation function and spectral density of such a process are given by (4.217) and (4.212), respectively. Applying these formulas to (7.39), i.e., setting

$$\beta = \frac{\epsilon}{2}, \qquad K = \epsilon^2 K_0 ,$$

we find that

$$\begin{aligned} &\langle y_1 y_{1\tau} \rangle = \langle y_2 y_{2\tau} \rangle = r(\tau) = \epsilon K_0 e^{-\epsilon|\tau|/2} , \\ &\langle y_1 y_{2\tau} \rangle = s(\tau) = 0 , \\ &\rho(\Omega) = \frac{\epsilon^2 K_0}{\Omega^2 + (\epsilon^2/4)}, \qquad \sigma(\Omega) = 0 . \end{aligned} \tag{7.43}$$

Moreover, according to (7.4) and (7.8), we have

$$\begin{aligned} \langle x x_\tau \rangle &= \epsilon K_0 e^{-\epsilon|\tau|/2} \cos \omega_0 \tau , \\ S[x;\omega] &= \frac{\epsilon^2 K_0}{(\omega - \omega_0)^2 + (\epsilon^2/4)} \qquad (\omega > 0) . \end{aligned} \tag{7.44}$$

Of course, these formulas could have been obtained directly from equation (7.35). The approach used here is convenient since it can be generalized to the case where the original equation is nonlinear. Such nonlinear equations will be studied in Vol. II, Part 2, which is devoted to the statistical theory of nonlinear oscillations, and methods will be given there for constructing higher-order approximations in the parameter ϵ. In the generalization to the nonlinear case, the slowly varying processes $y_1(t)$ and $y_2(t)$ remain Markovian, a fact which is very important, since this ensures that effective methods for analyzing the processes are available.

★ Everything just said can also be generalized to the case of equations of higher order, both linear and nonlinear. To illustrate how to go about making this generalization, we consider the fourth-order differential equation

$$(p^2 + \epsilon p + 1 + \epsilon\Delta)(p^2 + \epsilon p + 1 - \epsilon\Delta)\, x = \epsilon^2 \xi(t) , \tag{7.45}$$

where ϵ and Δ are constants, and $p = d/dt$. This equation describes the behavior of two "weakly detuned" resonant circuits in series, and has the expanded form

$$\left(\frac{d^2}{dt^2} + 1\right)(\ddot{x} + x) + 2\epsilon(\dddot{x} + \dot{x}) + \epsilon^2(\ddot{x} - \Delta^2 x) = \epsilon^2 \xi(t) .$$

Generalizing (7.36), we introduce four slowly varying processes:

$$\begin{aligned} y_1 &= x \cos t - \dot{x} \sin t , \\ y_2 &= -x \sin t - \dot{x} \cos t , \\ y_3 &= \frac{\ddot{x} + x}{\epsilon} \cos t - \frac{\dddot{x} + \dot{x}}{\epsilon} \sin t , \\ y_4 &= -\frac{\ddot{x} + x}{\epsilon} \sin t - \frac{\dddot{x} + \dot{x}}{\epsilon} \cos t . \end{aligned} \tag{7.46}$$

The inverse of the system (7.47) is

$$\begin{aligned} x &= y_1 \cos t - y_2 \sin t\,, \\ -\dot{x} &= y_1 \sin t + y_2 \cos t\,, \\ \frac{\ddot{x} + x}{\epsilon} &= y_3 \cos t - y_4 \sin t\,, \\ -\frac{\dddot{x} + \dot{x}}{\epsilon} &= y_3 \sin t + y_4 \cos t\,. \end{aligned} \tag{7.47}$$

Differentiating (7.46), we obtain the relations

$$\begin{aligned} \dot{y}_1 &= -(\ddot{x} + x) \sin t\,, \\ \dot{y}_2 &= -(\ddot{x} + x) \cos t\,, \\ \dot{y}_3 &= -\left[\left(\frac{d^2}{dt^2} + 1\right) \frac{\ddot{x} + x}{\epsilon}\right] \sin t\,, \\ \dot{y}_4 &= -\left[\left(\frac{d^2}{dt^2} + 1\right) \frac{\ddot{x} + x}{\epsilon}\right] \cos t\,. \end{aligned} \tag{7.48}$$

It follows from (7.45) and (7.48) that

$$\begin{aligned} \dot{y}_3 &= -\epsilon\left[-2\frac{\dddot{x} + \dot{x}}{\epsilon} + (1 + \Delta^2)\, x - \ddot{x} - x + \xi\right] \sin t\,, \\ \dot{y}_4 &= -\epsilon\left[-2\frac{\dddot{x} + \dot{x}}{\epsilon} + (1 + \Delta^2)\, x - \ddot{x} - x + \xi\right] \cos t\,. \end{aligned} \tag{7.49}$$

Of course, we must next express the right-hand sides of (7.49) in terms of y_1, y_2, y_3 and y_4, by substituting from (7.47). As a result, we obtain the following exact equations in "standard form":

$$\begin{aligned} \dot{y}_1 &= -\epsilon(y_3 \cos t - y_4 \sin t) \sin t\,, \\ \dot{y}_2 &= -\epsilon(y_3 \cos t - y_4 \sin t) \cos t\,, \\ \dot{y}_3 &= -\epsilon[2(y_3 \sin t + y_4 \cos t) + (1 + \Delta^2)(y_1 \cos t - y_2 \sin t) \\ &\quad -\epsilon(y_3 \cos t - y_4 \sin t) + \xi] \sin t\,, \\ \dot{y}_4 &= -\epsilon[2(y_3 \sin t + y_4 \cos t) + (1 + \Delta^2)(y_1 \cos t - y_2 \sin t) \\ &\quad -\epsilon(y_3 \cos t - y_4 \sin t) + \xi] \cos t\,. \end{aligned} \tag{7.50}$$

Then, in the right-hand sides of (7.50), we have to go over to sums of sines and cosines at the zero and double frequencies. In the first approximation, we can neglect the double-frequency "oscillatory terms," as well as nonoscillatory terms of higher order, which are proportional to ϵ^2. Thus, in the first approximation, we finally obtain the following system of equations:

$$\begin{aligned}
\dot{y}_1 &= \frac{\epsilon}{2} y_4 , \\
\dot{y}_2 &= -\frac{\epsilon}{2} y_3 , \\
\dot{y}_3 &= -\epsilon y_3 + \frac{\epsilon}{2}(1 + \Delta^2) y_2 - \epsilon \xi \sin t , \\
\dot{y}_4 &= -\epsilon y_4 - \frac{\epsilon}{2}(1 + \Delta^2) y_1 - \epsilon \xi \cos t .
\end{aligned} \tag{7.51}$$

The equations (7.51) are of the form (4.185), and for small ϵ, the processes y_1, y_2, y_3, y_4 can be regarded as Markov processes. This is equivalent to replacing the noise terms $\xi \sin t$ and $\xi \cos t$ in (7.51) by delta-correlated processes $\xi_1(t)$ and $\xi_2(t)$ with the properties (7.40) which go over into (7.41) after discarding oscillatory terms. The resulting system of equations separates into two completely independent systems. In fact, if we write $y_5 = -y_3$, these two systems

$$\begin{aligned}
\frac{\dot{y}_1}{\epsilon} &= \frac{y_4}{2} , \\
\frac{\dot{y}_4}{\epsilon} &= -y_4 - \frac{1 + \Delta^2}{2} y_1 - \xi_2 ,
\end{aligned} \tag{7.52}$$

and

$$\begin{aligned}
\frac{\dot{y}_2}{\epsilon} &= \frac{y_5}{2} , \\
\frac{\dot{y}_5}{\epsilon} &= -y_5 - \frac{1 + \Delta^2}{2} y_2 - \xi_1
\end{aligned} \tag{7.53}$$

are entirely identical.

From (7.52) and (7.53), it is easily deduced that each of the processes $y_1(t)$ and $y_2(t)$ has the spectral density

$$\rho(\omega) = \frac{K_0}{[2(\omega^2/\epsilon^2) - (1 + \Delta^2)/2]^2 + 4(\omega^2/\epsilon^2)}, \qquad \sigma(\omega) = 0.$$

Then, according to entry 5 of Table 1, p. 25, we have

$$r(\tau) = \frac{\epsilon}{2} \frac{K_0}{1 + \Delta^2} e^{-\epsilon|\tau|/2} \left(\cos \frac{\epsilon\Delta}{2} \tau + \frac{1}{\Delta} \sin \frac{\epsilon\Delta}{2} |\tau|\right). \tag{7.54}$$

In the case where the equation (7.45) is nonlinear, the equations for the slowly varying processes y_1, y_2, y_3, y_4 will also be nonlinear. However, the apparatus of Markov process theory can still be applied to these processes.★

3. The Amplitude and Phase of a Narrow-Band Process. Rayleigh Processes

The slowly varying processes $y_1(t)$ and $y_2(t)$ considered above are convenient because they are linearly related to the original narrow-band process $x(t)$. However, instead of $y_1(t)$ and $y_2(t)$, it is sometimes appropriate to consider the *amplitude* $A(t)$ and the *phase* $\varphi(t)$ of the process $x(t)$. These slowly varying processes have an obvious intuitive meaning, and are defined by the relations

$$A = \sqrt{y_1^2 + y_2^2}, \qquad \varphi = \arctan \frac{y_2}{y_1} \tag{7.55}$$

with inverses

$$y_1 = A \cos \varphi, \qquad y_2 = A \sin \varphi. \tag{7.56}$$

The processes (7.1), (7.17), (7.10) and (7.14) considered previously can be expressed in terms of the amplitude and phase as follows:

$$\begin{gathered} x = A \cos(\omega_0 t + \varphi), \qquad \tilde{x} = A \sin(\omega_0 t + \varphi), \\ z = Ae^{i\varphi}, \qquad \tilde{z} = Ae^{i\omega_0 t + i\varphi}. \end{gathered} \tag{7.57}$$

The amplitude and phase are related to $x(t)$ in a nonlinear fashion. Therefore, to calculate their statistical characteristics, we have to know the probability distribution of the process $x(t)$. For simplicity, we shall only consider the case where $x(t)$ is assumed to be a Gaussian process. Then, knowledge of the correlation function (7.4) is sufficient to determine all possible distributions not only of $x(t)$ itself, but also of the slowly varying processes $y_1(t)$ and $y_2(t)$. In fact, $y_1(t)$ and $y_2(t)$ are also Gaussian with zero means, since they can be expressed linearly in terms of $x(t)$. Once we know the correlation functions $\langle y_1 y_{1\tau}\rangle = \langle y_2 y_{2\tau}\rangle = r(\tau)$ and the cross-correlation function $\langle y_1 y_{2\tau}\rangle = s(\tau)$, we can find the various probability distributions of $y_1(t)$ and $y_2(t)$.

Since $s(\tau)$ is odd, the cross correlation between the values of $y_1(t)$ and $y_2(t)$, taken at the same time t, equals zero, and hence

$$w(y_1, y_2) = \frac{1}{2\pi\sigma^2} e^{-(y_1^2+y_2^2)/2\sigma^2}, \tag{7.58}$$

where $\sigma^2 = r(0)$. Making the change of variables (7.56) and integrating over all possible values of the phase from 0 to 2π, we find that the amplitude has the *Rayleigh distribution*

$$w(A) = \frac{A}{\sigma^2} e^{-A^2/2\sigma^2}. \tag{7.59}$$

A random process $A(t)$ which is the amplitude of a narrow-band Gaussian process is called a *Rayleigh process.*

To find the two-dimensional probability density $w(A, A_\tau)$ of the amplitude, we have to use the distribution of the random variables $y_1, y_2, y_{1\tau}, y_{2\tau}$, which can be found from their correlation matrix

$$\begin{Vmatrix} r(0) & 0 & r(\tau) & s(\tau) \\ 0 & r(0) & -s(\tau) & r(\tau) \\ r(\tau) & -s(\tau) & r(0) & 0 \\ s(\tau) & r(\tau) & 0 & r(0) \end{Vmatrix}.$$

Then we go from $y_1, y_2, y_{1\tau}, y_{2\tau}$ to A, A_τ, φ, φ_τ, and integrate over all values of φ and φ_τ. As is well known,[1] the result is

$$w(A, A_\tau) = \frac{AA_\tau}{\sigma^4(1-Q^2)} I_0\left(\frac{Q}{1-Q^2}\frac{AA_\tau}{\sigma^2}\right) \exp\left\{-\frac{A^2+A_\tau^2}{2\sigma^2(1-Q^2)}\right\}, \quad (7.60)$$

where

$$Q = \frac{\sqrt{r^2(\tau)+s^2(\tau)}}{r(0)}, \qquad \sigma^2 = r(0),$$

and $I_0(z) = J_0(iz)$ is the modified Bessel function of the first kind, of order zero.

In the case of a narrow-band noise process described by a differential equation, the behavior of the amplitude and phase is also described by differential equations, just as in the case of the functions y_1 and y_2. For example, consider the process described by the equation (7.35). Differentiating (7.55), we find that

$$\dot{A} = \frac{y_1\dot{y}_1 + y_2\dot{y}_2}{A}, \qquad \dot{\varphi} = \frac{y_1\dot{y}_2 - y_2\dot{y}_1}{A^2}. \quad (7.61)$$

Substituting (7.39) into (7.61) and using (7.56), to the first approximation we obtain the equations

$$\begin{aligned} \dot{A} &= -\frac{\epsilon}{2}A - \epsilon\xi_1\cos\varphi - \epsilon\xi_2\sin\varphi \equiv \epsilon F_1, \\ \dot{\varphi} &= \frac{\epsilon}{A}[\xi_1\sin\varphi - \xi_2\cos\varphi] \equiv \epsilon F_2, \end{aligned} \quad (7.62)$$

where $\xi_1(t)$ and $\xi_2(t)$ are Gaussian processes with the properties (7.41).

To write down the Fokker-Planck equation for the amplitude and phase, which vary in the way described by (7.62), we apply formula (4.194). First, we calculate the coefficients appearing in (4.194).

[1] See e.g., W. B. Davenport, Jr. and W. L. Root, *An Introduction to the Theory of Random Signals and Noise*, McGraw-Hill Book Co., Inc., N.Y. (1958), p. 162.

It is easily seen that

$$\begin{aligned}\int_{-\infty}^{0} \mathbf{K}[F_1, F_{1\tau}]\, d\tau &= \int_{-\infty}^{0} \langle (\xi_1 \cos\varphi + \xi_2 \sin\varphi)\,(\xi_{1\tau} \cos\varphi + \xi_{2\tau} \sin\varphi)\rangle\, d\tau \\ &= \cos^2\varphi \int_{-\infty}^{0} \langle \xi_1 \xi_{1\tau}\rangle\, d\tau + \sin^2\varphi \int_{-\infty}^{0} \langle \xi_2 \xi_{2\tau}\rangle\, d\tau \\ &= \frac{1}{2} K_0\,,\end{aligned} \tag{7.63}$$

and similarly,

$$\begin{aligned}&\int_{-\infty}^{0} \mathbf{K}[F_2, F_{2\tau}]\, d\tau = \frac{1}{2A^2} K_0\,, \\ &\int_{-\infty}^{0} \mathbf{K}[F_1, F_{2\tau}]\, d\tau = 0\,, \\ &\int_{-\infty}^{0} \mathbf{K}\left[\frac{\partial F_1}{\partial \varphi}, F_{2\tau}\right] d\tau = \frac{1}{A} \int_{-\infty}^{0} \langle (\xi_1 \sin\varphi - \xi_2 \cos\varphi) \\ &\qquad\qquad \times (\xi_{1\tau} \sin\varphi - \xi_{2\tau} \cos\varphi)\rangle\, d\tau \\ &\qquad = \frac{1}{2A} K_0\,.\end{aligned} \tag{7.64}$$

Thus, according to (4.194), the required Fokker-Planck equation has the form

$$\dot{w}(A, \varphi) = -\epsilon \frac{\partial}{\partial A}\left[\left(-\frac{A}{2} + \frac{\epsilon}{2A} K_0\right) w\right] + \frac{\epsilon^2}{2} K_0 \left[\frac{\partial^2 w}{\partial A^2} + \frac{1}{A^2} \frac{\partial^2 w}{\partial \varphi^2}\right], \tag{7.65}$$

which corresponds to the following fluctuation equations for the amplitude and phase:

$$\begin{aligned}\dot{A} &= -\frac{\epsilon}{2} A + \frac{\epsilon^2}{2A} K_0 + \epsilon \xi_1\,, \\ \dot{\varphi} &= \epsilon \frac{\xi_2}{A}\,.\end{aligned} \tag{7.66}$$

Here, ξ_1 and ξ_2 are the same kind of independent random processes as in equations (7.39) and (7.62), with the same properties (7.41).

Although the term $\epsilon^2 K_0/2A$ is of order ϵ^2, we have retained it in (7.66), because the intensity coefficient K_0 of the fluctuations may be quite large. It is best to assume that K_0 is of order $1/\epsilon$, and write

$$K_0 = \frac{\kappa}{\epsilon}.$$

Then all the terms in (7.65) are of the first order in ϵ, and it is convenient to write the noise terms in (7.66) as $\sqrt{\epsilon}\eta_1$ and $\sqrt{\epsilon}\eta_2$. As a result, the first of the equations (7.66), which describes the behavior of the amplitude, becomes

$$\dot{A} = -\frac{\epsilon}{2}\left(A - \frac{\kappa}{A}\right) + \sqrt{\epsilon}\,\eta_1, \tag{7.67}$$

where

$$\langle \eta_1 \rangle = 0, \qquad \langle \eta_1 \eta_{1\tau} \rangle = \kappa\delta(\tau).$$

The one-dimensional Fokker-Planck equation corresponding to (7.67) is

$$\frac{2}{\epsilon}\dot{w}(A) = \frac{\partial}{\partial A}\left[\left(A - \frac{\kappa}{A}\right)w\right] + \kappa\frac{\partial^2 w}{\partial A^2}. \tag{7.68}$$

We already know the form of the stationary solution (4.49) of the arbitrary one-dimensional Fokker-Planck equation (4.33). Applying (4.49) to (7.68), we find that the stationary amplitude distribution is

$$w(A) = Ce^{-(A^2/2\kappa)+\ln A} = \frac{A}{\kappa}e^{-A^2/2\kappa}, \tag{7.69}$$

which, of course, is the same as (7.59), with $\kappa = \sigma^2$.

Since the changes in amplitude represent a Markov process, the expression for the transition probability $p(A_\tau, A)$ is of great interest, since all possible probability distributions, in particular, the two-dimensional distribution

$$w(A, A_\tau) = p(A_\tau, A)\,w(A),$$

can be expressed in terms of $p(A_\tau, A)$. Using (7.60) and (7.59), we obtain the following expression for the transition probability

$$\begin{aligned} p(A_\tau, A) &= \frac{w(A, A_\tau)}{w(A)} \\ &= \frac{A_\tau}{\sigma^2(1-Q^2)} I_0\left(\frac{Q}{1-Q^2}\frac{AA_\tau}{\sigma^2}\right)\exp\left\{-\frac{A_\tau^2 + Q^2A^2}{2\sigma^2(1-Q^2)}\right\}. \end{aligned} \tag{7.70}$$

Because of the relations (7.43), which were found earlier for the case described by the equation (7.35), the coefficient $Q(\tau)$ in (7.70) has the form

$$Q(\tau) = e^{-\epsilon|\tau|/2}. \tag{7.71}$$

If $Q(\tau)$ has a different form, $A(t)$ will not be a Markov process, and the conditional probability (7.70) will not play such an important role as in the case of a Markov process. Substituting (7.71) into (7.70), and replacing σ^2 by κ, we find that the nonstationary amplitude distribution is

$$w(A) = \frac{A}{\kappa(1 - e^{-\epsilon\tau})} I_0\left(\frac{AA_0}{2\kappa \sinh(\epsilon\tau/2)}\right) \exp\left\{-\frac{A^2}{2\kappa(1 - e^{-\epsilon\tau})} - \frac{A_0^2}{2\kappa(e^{\epsilon\tau} - 1)}\right\}, \tag{7.72}$$

whose time variation is described by (7.68) if the initial distribution is a delta function, i.e., if

$$w(A) = \delta(A - A_0) \quad \text{for} \quad \tau = 0.$$

This is one of several examples where a Fokker-Planck equation with a nonlinear "average velocity" K_1 (cf. p. 58) has an exact nonstationary solution, in the form of an analytic expression involving elementary functions.

It is sometimes required to find the statistical characteristics of a narrow-band random process $x(t)$ or of the other processes (7.56) and (7.57), given the statistical characteristics of the processes describing the variation of amplitude and phase. In its general form, this is a complicated problem, since the transformation from $A(t)$ and $\varphi(t)$ to the processes mentioned is nonlinear. However, the problem can be considerably simplified in the case where the amplitude and phase processes are statistically independent of each other. We now show how to calculate the correlation functions of the processes

$$z(t) = Ae^{i\varphi}, \qquad x(t) = A\cos(\omega_0 t + \varphi)$$

in this case. Using the independence of the amplitude and phase, we can write the correlation function

$$\langle zz_\tau^* \rangle = \langle AA_\tau\, e^{i(\varphi-\varphi_\tau)} \rangle \tag{7.73}$$

as

$$\langle zz_\tau^* \rangle = \langle AA_\tau \rangle \langle e^{i(\varphi-\varphi_\tau)} \rangle .$$

The average $\langle e^{i(\varphi-\varphi_\tau)} \rangle$ is the value at $u = -1$ of the characteristic function

$$\Theta_\tau(u) = \langle e^{iu(\varphi_\tau-\varphi)} \rangle \tag{7.74}$$

of the phase increment $\varphi_\tau - \varphi$. Therefore, we have

$$\langle zz_\tau^* \rangle = \langle AA_\tau \rangle\, \Theta_\tau(-1) . \tag{7.75}$$

Then, using (7.16) we find that the correlation function of the real random process $x(t)$ is

$$\langle xx_\tau \rangle = \tfrac{1}{2} \langle AA_\tau \rangle \operatorname{Re} [\Theta_\tau(-1)\, e^{-i\omega_0\tau}] . \tag{7.76}$$

When the phase increment has a symmetric distribution about zero, the characteristic function $\Theta(u)$ is real and (7.76) becomes

$$\langle xx_\tau \rangle = \tfrac{1}{2} \langle AA_\tau \rangle\, \Theta_\tau(-1) \cos \omega_0\tau . \tag{7.77}$$

According to (7.4), this means that

$$r(\tau) = \tfrac{1}{2} \langle AA_\tau \rangle\, \Theta_\tau(-1) , \qquad s(\tau) = 0 . \tag{7.78}$$

As an example, consider the case where the phase changes represent a Wiener process (p. 135), i.e., a "pure diffusion process." Then $\varphi_\tau - \varphi$ is a normally distributed random variable, whose mean value is zero and whose variance grows linearly with time:

$$\mathrm{D}(\varphi_\tau - \varphi) = K_\varphi |\tau| ,$$

In this case,

$$\Theta_\tau(u) = \exp \{-\tfrac{1}{2} K_\varphi |\tau| u^2\}$$

and formula (7.77) becomes

$$\langle xx_\tau \rangle = \tfrac{1}{2} \langle AA_\tau \rangle\, e^{-K_\varphi|\tau|/2} \cos \omega_0\tau . \tag{7.79}$$

4. Quasi-Rayleigh Processes

Suppose we are given a narrow-band Gaussian process $\xi(t)$ with correlation function $\sigma^2 R(\tau)$, and suppose we add to $\xi(t)$ a "sine wave" $E\cos(\omega_0 t + \vartheta_0)$, with fixed amplitude E and random initial phase ϑ_0 which is statistically independent of $\xi(t)$. Then, the resulting process

$$x(t) = \xi(t) + E\cos(\omega_0 t + \vartheta_0) \tag{7.80}$$

is no longer Gaussian. For (7.80) to be a stationary process, it is necessary that the initial phase ϑ_0 be "completely random," i.e., that it have the uniform distribution law

$$w(\vartheta_0) = \frac{1}{2\pi}. \tag{7.81}$$

Then, using the statistical independence of $\xi(t)$ and $E\cos(\omega_0 t + \vartheta_0)$, we can easily find the correlation function and characteristic functions of $x(t)$:

$$\begin{aligned}
\langle x x_\tau \rangle &= \sigma^2 R(\tau) + \frac{E^2}{2}\cos\omega_0\tau\,, \\
\langle e^{iux} \rangle &= e^{-\sigma^2 u^2/2} J_0(uE)\,, \\
\langle e^{iux + iu_\tau x_\tau} \rangle &= \exp\left\{-\frac{\sigma^2}{2}(u^2 + 2uu_\tau R(\tau) + u_\tau^2)\right\} \\
&\quad \times J_0\left(E\sqrt{u^2 + u_\tau^2 + 2uu_\tau\cos\omega_0\tau}\right),
\end{aligned} \tag{7.82}$$

where $J_0(z)$ is the Bessel function of the first kind, of order zero.

The narrow-band process $x(t)$ can also be written as an oscillation of the form

$$x(t) = B(t)\cos[\omega_0 t + \psi(t)] \tag{7.83}$$

with slowly varying amplitude B and phase ψ, which we denote by new letters to distinguish them from the amplitude A and phase φ of the Gaussian process discussed in the previous section. We now find the probability density $w(B)$. First, we write the process $\xi(t)$ as

$$\xi = y_1\cos\omega_0 t - y_2\sin\omega_0 t\,, \tag{7.84}$$

where y_1 and y_2 are random variables with the normal distribution law (7.58). Substituting (7.84) into (7.80), and comparing the result with (7.83), we have

$$y_1 + E\cos\vartheta_0 = B\cos\psi\,, \quad y_2 + E\sin\vartheta_0 = B\sin\psi\,, \tag{7.85}$$

so that

$$y_1^2 + y_2^2 = B^2 + E^2 - 2BE\cos(\psi - \vartheta_0)\,. \tag{7.86}$$

By making the change of variables (7.85), we go from the distribution

$$w(y_1, y_2, \vartheta_0) = \frac{1}{(2\pi\sigma)^2}\exp\left\{-\frac{y_1^2 + y_2^2}{2\sigma^2}\right\} \tag{7.87}$$

to $w(B, \psi, \vartheta_0)$. Bearing in mind that

$$dy_1 dy_2 d\vartheta_0 = BdBd\psi d\vartheta_0\,,$$

and substituting (7.86) into (7.87), we find that

$$w(B, \psi, \vartheta_0) = \frac{B}{(2\pi\sigma)^2}\exp\left\{-\frac{1}{2\sigma^2}[B^2 + E^2 - 2BE\cos(\psi - \vartheta_0)]\right\}.$$

Then, integrating this expression with respect to ψ and ϑ_0, we obtain the amplitude distribution

$$w(B) = \frac{B}{\sigma^2}\exp\left\{-\frac{B^2 + E^2}{2\sigma^2}\right\} I_0\left(\frac{EB}{\sigma^2}\right), \tag{7.88}$$

which is a generalization of (7.59). A calculation of the mean value

$$\langle B\rangle = \int_0^\infty w(B)\, BdB$$

leads to the following expression:

$$\langle B\rangle = \sqrt{\frac{\pi}{2}}\,\sigma\, e^{-E^2/4\sigma^2}\left[\left(1 + \frac{E^2}{2\sigma^2}\right) I_0\left(\frac{E^2}{4\sigma^2}\right) + \frac{E^2}{2\sigma^2} I_1\left(\frac{E^2}{4\sigma^2}\right)\right]. \tag{7.89}$$

The random process $B(t)$ will be called a *quasi-Rayleigh process*. Its importance in radio engineering is due to the fact that a received radio signal often contains both noise and a useful signal which has

the form of a sine wave whose amplitude can be regarded as constant, for simplicity.

There is a different way in which one can arrive at a quasi-Rayleigh process. Suppose that in formula (7.1), $y_1(t)$ and $y_2(t)$ are stationary Gaussian processes whose mean values are not zero. Then the process $x(t)$ will be nonstationary. However, stationarity can be achieved by making the time origin (i.e., the time to which all other times are referred) indeterminate; in fact, it is sufficient to introduce a physically imperceptible error in the location of the time origin, with order of magnitude equal to the high-frequency period $2\pi/\omega_0$. Then the amplitude of the process (7.1) will be a quasi-Rayleigh process with

$$E = \sqrt{\langle y_1 \rangle^2 + \langle y_2 \rangle^2}\,.$$

The quasi-Rayleigh process $B(t)$ is not a Markov process, even when $y_1(t)$ and $y_2(t)$ are exponentially correlated random processes. Instead, $B(t)$ is a component of the two-dimensional Markov process $(B(t), \psi(t))$ decribed by the equations

$$\begin{aligned} \frac{\dot{B}}{\epsilon} &= -\frac{B}{2} + \frac{\sigma^2}{2B} + \frac{E}{2}\cos\chi + \xi_1\,, \\ \frac{\dot{\chi}}{\epsilon} &= -\frac{E}{2B}\sin\chi + \frac{\xi_2}{B} \qquad (\chi = \psi - \vartheta_0)\,, \end{aligned} \tag{7.90}$$

which generalizes (7.66). Here, ξ_1 and ξ_2 are independent Gaussian delta-correlated processes such that

$$\langle \xi_1 \rangle = \langle \xi_2 \rangle = 0\,, \qquad \langle \xi_1 \xi_{1\tau} \rangle = \langle \xi_2 \xi_{2\tau} \rangle = \frac{\sigma^2}{\epsilon}\,\delta(\tau)\,.$$

However, for simplitiy, in solving certain problems it turns out to be convenient to replace the quasi-Rayleigh process $B(t)$ by the one-dimensional Markov process corresponding to the equation

$$\frac{\dot{B}}{\epsilon} = -\frac{B}{2} + \frac{\sigma^2}{2B} + \frac{E}{2}\,\frac{I_1(EB/\sigma^2)}{I_0(EB/\sigma^2)} + \xi_1\,, \tag{7.91}$$

which has the same stationary distribution as (7.88). This replacement is satisfactory both for large and small values of E/σ.

PART 2

Nonlinear Transformations of Signals and Noise

CHAPTER 8

Zero-Memory Nonlinear Transformations

In studying nonlinear transformations of random signals, just as in the case of nonlinear oscillations in general, it is difficult to develop a universal practical method of solving problems. Instead, different problems require the use of different methods, which are justified in each case on physical grounds. It is appropriate to distinguish two basic cases, depending on the character of the nonlinear system in question:

1. The simplest nonlinear systems are those for which the value of the output function $\eta(t)$ at any instant of time is determined only by the value of the input function $\xi(t)$ at the same instant of time, i.e.,

$$\eta(t) = g[\xi(t)], \tag{8.1}$$

where $g(\xi)$ is a nonlinear function (see Figure 4a). Such a nonlinear transformation is said to have *zero memory* (or to have *no inertia*). A closely related group of nonlinear transformations are those such that the input function $\xi(t)$ is subjected to an additional transformation by a linear system S_1, and the output function is subjected to an additional transformation by a linear system S_2, where the action of the systems S_1 and S_2 is not affected by the presence of the nonlinear device (see Figure 4b). Such a transformation can be written in the form

$$\eta(t) = L_2 g[L_1 \xi(t)], \tag{8.2}$$

where L_1 and L_2 are the linear operators describing the behavior of

the systems S_1 and S_2. We know how the correlation functions and quasi-moment functions of random signals transform when the signals undergo linear transformations. Therefore, the study of transformations of the type (8.1) is equivalent to the study of transformations of the type (8.2).

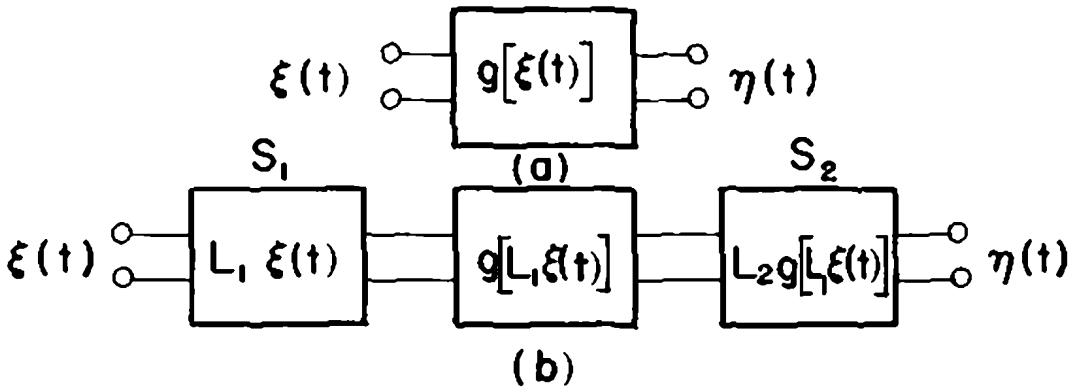

Fig. 4. Examples of nonlinear zero-memory transformations.

2. A more general and complicated class of transformations consists of nonlinear transformations with *memory* (or *inertia*), in particular, transformations described by nonlinear differential equations. In Figure 5, we show one of the simplest examples of such a transformation, i.e., a system whose behavior is described by a first-order nonlinear differential equation. Let $\xi(t)$ and $\eta(t)$ be the random input and output voltages, respectively, and let D be a nonlinear element (a diode) whose current-voltage characteristic is

$$I = F(V),$$

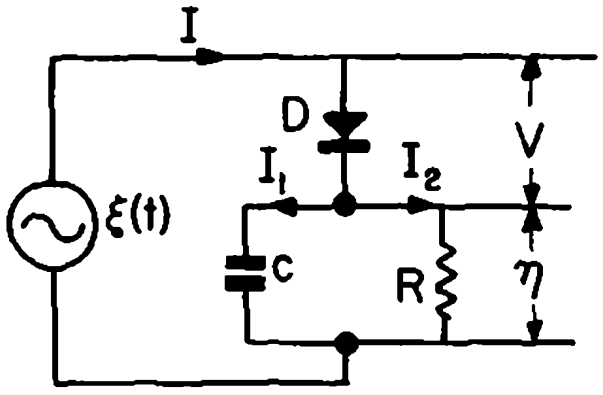

Fig. 5. The diode detector, an example of a nonlinear transformation with memory.

where $V = \xi - \eta$. Assuming that the internal resistance of the input-signal generator is zero, and using the obvious relations

$$I_1 + I_2 = I, \qquad \eta = \frac{1}{C}\int I_1\, dt = I_2 R,$$

we obtain the differential equation

$$\frac{d\eta}{dt} + \frac{1}{RC}\eta = \frac{1}{C}F(\xi - \eta). \tag{8.3}$$

It is difficult to solve this nonlinear equation in the general case where the function F is arbitrary, even for a deterministic input ξ, since the solution cannot be expressed in terms of a finite number of integrals. Therefore, it is natural that in the case of random functions, the solution of (8.3) involves great difficulties.

We now make a few remarks about the form of the function $g(\xi)$ appearing in (8.1) and (8.2). This function is called the *transfer* (or *dynamic*) *characteristic* of the nonlinear device, and is usually determined experimentally from the current-voltage characteristics of nonlinear elements like vacuum tubes. Then, to facilitate subsequent calculations, one makes some suitable approximation to $g(\xi)$. Three approximation methods which are convenient for theoretical purposes are to approximate $g(\xi)$ by 1) a *polynomial*, 2) a *broken line* (the so-called *piecewise linear approximation*), and 3) by an *exponential*. Each of these methods has its advantages and disadvantages. It should be kept in mind that the requirement that the approximation be accurate and the requirement that the resulting analytic expressions be simple are in a certain sense antithetical, and as a rule incompatible.

In the next two chapters, we shall analyze what happens to random signals when they are acted upon by nonlinear systems. Our problem can be formulated as follows: *Given the parameters of the system and the statistical characteristics of the input signal* $\xi(t)$, *find the statistical characteristics of the output signal* $\eta(t)$.[1]

[1] Cf. P. I. Kuznetsov, R. L. Stratonovich and V. I. Tikhonov, *Passage of some random functions through linear systems*, Avtomat. i Telemekh., **14**, 144 (1953); *Passage of random functions through nonlinear systems*, ibid., **14**, 375 (1953).

1. Transformations of Probability Densities

For nonlinear zero-memory transformations of the form (8.1), the basic solution of the problem just formulated is a consequence of the following familiar result: Suppose we know the n-dimensional probability density $w_\xi(\xi_1, \ldots, \xi_n)$ of the random variables

$$\xi_1 = \xi(t_1), \ldots, \xi_n = \xi(t_n),$$

and suppose we want to find the probability density $w_\eta(\eta_1, \ldots, \eta_n)$ of the new random variables

$$\begin{aligned} \eta_1 &= g_1(\xi_1, \ldots, \xi_n), \\ &\cdots \\ \eta_n &= g_n(\xi_1, \ldots, \xi_n), \end{aligned} \tag{8.4}$$

where the functions $g_1, \ldots, g_n$ are piecewise continuous. If the inverse of the system (8.4) is

$$\begin{aligned} \xi_1 &= h_1(\eta_1, \ldots, \eta_n), \\ &\cdots \\ \xi_n &= h_n(\eta_1, \ldots, \eta_n), \end{aligned} \tag{8.5}$$

where the functions $h_1, \ldots, h_n$ are single-valued, then the probability density w is given by the formula

$$w_\eta(\eta_1, \ldots, \eta_n) = w_\xi[h_1(\eta_1, \ldots, \eta_n), \ldots, h_n(\eta_1, \ldots, \eta_n)] \mid D_n \mid, \tag{8.6}$$

where D_n is the Jacobian of the transformation from the random variables $\xi_1, \ldots, \xi_n$ to the random variables $\eta_1, \ldots, \eta_n$, i.e.,

$$D_n = \begin{vmatrix} \dfrac{\partial h_1}{\partial \eta_1} & \cdots & \dfrac{\partial h_1}{\partial \eta_n} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial h_n}{\partial \eta_1} & \cdots & \dfrac{\partial h_n}{\partial \eta_n} \end{vmatrix}. \tag{8.7}$$

In cases where the inverse functions $h_1, \ldots, h_n$ are not single-valued, we have to sum the right hand side of (8.6) over each of the subregions involved. For the one-dimensional case, this

transformation rule was derived on pp. 9-10. Applying (8.6) to the case of transformations of the form (8.6), we obtain

$$w_\eta(\eta_1, \eta_2) = w_\xi[h(\eta_1), h(\eta_2)] \mid h'(\eta_1)\, h'(\eta_2) \mid \tag{8.8}$$

for $n = 2$, where $\xi = h(\eta)$ is the inverse of the function $g(\xi)$.

Difficulties can arise in applying formulas (8.6) and (8.8) to the nonlinear zero-memory transformations of practical interest. For example, if the transfer characteristic $g(\xi)$ in (8.1) is a polynomial of degree higher than three, then it is hard to find the inverse function $h(\eta)$ in general form, i.e., it is hard to solve (8.1) for ξ analytically. Thus, in this case, it is appropriate to approximate the function $g(\xi)$. If we make a piecewise linear approximation to $g(\xi)$, in many cases the function $h(\eta)$ turns out to be discontinuous with an infinite derivative $h'(\eta)$ at certain points. Then, at these points, the probability density of $\eta(t)$ will have delta-function discontinuities.

We now consider a few special cases. Let

$$\eta_1 = g_1(\xi_1) = \xi_1, \qquad \eta_2 = g_2(\xi_1, \xi_2), \tag{8.9}$$

where the inverse functions

$$\xi_1 = h_1(\eta_1) = \eta_1, \qquad \xi_2 = h_2(\eta_1, \eta_2) \tag{8.10}$$

are single-valued. The Jacobian of this transformation is

$$D_2 = \begin{vmatrix} 1 & 0 \\ \dfrac{\partial h_2}{\partial \eta_1} & \dfrac{\partial h_2}{\partial \eta_2} \end{vmatrix} = \frac{\partial h_2}{\partial \eta_2},$$

and therefore

$$w_\eta(\eta_1, \eta_2) = w_\xi[\eta_1, h_2(\eta_1, \eta_2)] \left| \frac{\partial h_2}{\partial \eta_2} \right|. \tag{8.11}$$

To obtain the one-dimensional probability density of the random variable η_2, we integrate (8.11) with respect to η_1:

$$w(\eta_2) = \int w_\xi[\eta_1, h_2(\eta_1, \eta_2)] \left| \frac{\partial h_2}{\partial \eta_2} \right| d\eta_1. \tag{8.12}$$

Using (8.12), we can derive the following expressions for the probability densities of the sum, difference, product and quotient of the two random variables ξ_1 and ξ_2:

$$w(\xi_1 + \xi_2) = w(\eta) = \int w_\xi(\xi_1, \eta - \xi_1)\, d\xi_1 , \tag{8.13}$$

$$w(\xi_1 - \xi_2) = w(\eta) = \int w_\xi(\xi_1, \xi_1 - \eta)\, d\xi_1 , \tag{8.14}$$

$$w(\xi_1\xi_2) = w(\eta) = \int w_\xi\left(\xi_1, \frac{\eta}{\xi_1}\right) \frac{d\xi_1}{|\xi_1|} , \tag{8.15}$$

$$w\left(\frac{\xi_2}{\xi_1}\right) = w(\eta) = \int w_\xi(\xi_1, \eta\xi_1)\, |\xi_1|\, d\xi_1 . \tag{8.16}$$

If the random variables ξ_1 and ξ_2 are independent, with probability densities $p(\xi_1)$ and $q(\xi_2)$, then in these expressions we have to set

$$w_\xi(\xi_1, \xi_2) = p(\xi_1)\, q(\xi_2) . \tag{8.17}$$

The next two examples illustrate the use of formulas (8.15) and (8.16).

1. Suppose we want to find the probability density of the product of two correlated random variables ξ_1 and ξ_2, whose joint probability density is normal, i.e.,

$$w_\xi(\xi_1, \xi_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1 - R^2}} \exp\left\{-\frac{1}{2(1 - R^2)}\left(\frac{\xi_1^2}{\sigma_1^2} + \frac{\xi_1^2}{\sigma_2^2} - \frac{2R\xi_1\xi_2}{\sigma_1\sigma_2}\right)\right\} .$$

Then, applying (8.15) with $\eta = \xi_1\xi_2$, we obtain

$$\begin{aligned} w(\eta) &= \frac{1}{\pi\sigma_1\sigma_2\sqrt{1 - R^2}} \exp\left\{\frac{R\eta}{\sigma_1\sigma_2(1 - R^2)}\right\} \\ &\quad \times \int_0^\infty \exp\left\{-\frac{1}{2\sigma_1^2(1 - R^2)}\left(\xi_1^2 + \frac{\sigma_1^2\eta^2}{\sigma_2^2\xi_1^2}\right)\right\} \frac{d\xi_1}{\xi_1} \qquad (8.18) \\ &= \frac{1}{\pi\sigma_1\sigma_2\sqrt{1 - R^2}} \exp\left\{\frac{R\eta}{\sigma_1\sigma_2(1 - R^2)}\right\} K_0\left(\frac{\eta}{\sigma_1\sigma_2(1 - R^2)}\right) , \end{aligned}$$

where $K_0(z) = (\pi i/2)H_0^{(1)}(iz)$ is the modified Bessel function of the second kind, of order zero.[2]

2. Next, we calculate the probability density of the ratio $\eta = \xi_2/\xi_1$ of two normally distributed, correlated random variables ξ_1 and ξ_2. According to (8.16), we have

$$w(\eta) = \frac{1}{\pi\sigma_1\sigma_2\sqrt{1-R^2}} \int_0^\infty \exp\left\{-\frac{\xi_1^2}{2(1-R^2)}\left(\frac{1}{\sigma_1^2} + \frac{\eta^2}{\sigma_2^2} - \frac{2R\eta}{\sigma_1\sigma_2}\right)\right\} \xi_1\, d\xi_1$$

$$= \frac{\sigma_1\sigma_2\sqrt{1-R^2}}{\pi(\sigma_2^2 - 2R\sigma_1\sigma_2 + \sigma_1^2\eta^2)}. \tag{8.19}$$

2. Moment Functions for Polynomial Transfer Characteristics

Let the transfer characteristic $\eta = g(\xi)$ of the nonlinear device be an analytic function in a certain interval, so that it can be expanded in a Taylor's series

$$g(\xi) = a_0 + a_1(\xi - c) + a_2(\xi - c)^2 + \dots, \tag{8.20}$$

where

$$a_k = \frac{1}{k!} g^{(k)}(c).$$

Then, depending on the required degree of approximation, we retain a certain number of terms in (8.22), say the first $n + 1$ terms:

$$g(\xi) = a_0 + a_1(\xi - c) + \dots + a_n(\xi - c)^n. \tag{8.21}$$

Even if this approach is not possible, we can still always find a polynomial which approximates the function $g(\xi)$ with the required accuracy for a given input process $\xi(t)$. In choosing this approximation, it is important that the polynomial be a good approximation to $g(\xi)$ on that part of the ξ-axis corresponding to values of $\xi(t)$ which have high probability. However, the approximation can be quite inaccurate for parts of the transfer characteristic corresponding to values of $\xi(t)$ which occur with relatively small probability.

[2] See I. M. Ryshik and I. S. Gradstein, *op. cit.*, formula (6.447.2), p. 318.

Another way of finding the best approximation to $g(\xi)$ is based on the method of least square error and goes as follows: If the approximating polynomial is to have $n + 1$ terms, we first form the difference

$$g(\xi) - \sum_{k=0}^{n} a_k(\xi - c)^k .$$

Then the unknown coefficients a_k are determined from the condition that the integral of the square of the error, i.e., of the quantity

$$I = \int_a^b \left[g(c + x) - \sum_{k=0}^{n} a_k x^k\right]^2 dx , \tag{8.22}$$

should take its minimum value. The limits of integration a and b are chosen in such a way that the value of the random variable ξ is overwhelmingly likely to fall in the interval $a + c \leqslant \xi \leqslant b + c$, which means that

$$\int_{a+c}^{b+c} w(\xi)\, d\xi \approx 1 .$$

If the input process $\xi(t)$ is specified by its moments, then without using these moments to reconstruct the probability density $w(\xi)$, it is convenient to set

$$c = m_1 , \qquad b = -a = \lambda \sqrt{m_2 - m_1^2} . \tag{8.23}$$

Moreover, if the probability density $w(\xi)$ does not differ too strongly from a normal distribution, we can choose λ equal to about 2 or 3. The necessary conditions

$$\frac{\partial I}{\partial a_0} = 0 , \ldots, \frac{\partial I}{\partial a_n} = 0$$

for the expression (8.22) to have a minimum lead to the system of equations

$$\sum_{j=0}^{n} \frac{1 - (-1)^{j+k+1}}{j + k + 1} b^{j+k+1} a_j = \int_a^b x^k g(x + c)\, dx , \tag{8.24}$$

from which the coefficients $a_0, a_1, \ldots, a_n$ can be determined.

Thus, for the rest of this section, we shall assume that instead of the general transformation (8.1), we are dealing with a transformation of the form

$$y(t) = a_1 x(t) + \dots + a_n x^n(t), \tag{8.25}$$

where

$$y(t) = \eta(t) - a_0, \qquad x(t) = \xi(t) - c = \xi(t) - \langle \xi \rangle.$$

To find the moment functions of $y(t)$, which will be denoted by $\tilde{m}_n$, we multiply the expression (8.25) by similar expressions corresponding to different instants of time. We then average the result, obtaining

$$\begin{aligned}
\tilde{m}_1(t) &= a_2 \mu_2(t, t) + a_3 \mu_3(t, t, t) + \dots + a_n \mu_n(t, \dots, t), \\
\tilde{m}_2(t_1, t_2) &= a_1^2 \mu_2(t_1, t_2) + a_1 a_2 [\mu_3(t_1, t_2, t_2) + \mu_3(t_1, t_1, t_2)] \\
&\quad + a_1 a_3 [\mu_4(t_1, t_2, t_2, t_2) + \mu_4(t_1, t_1, t_1, t_2] \\
&\quad + a_2^2 \mu_4(t_1, t_1, t_2, t_2) + \dots,
\end{aligned} \tag{8.26}$$

Here, the quantities

$$\mu_n(t_1, \dots, t_n) = \langle x(t_1) \dots x(t_n) \rangle$$

are the moment functions of the process $x(t)$, and are equal to the central moments of $\xi(t)$ when $c = \langle \xi(t) \rangle$.

It is apparent from the formulas (8.26) that the moment functions of the output process $\eta(t)$ can be expressed as linear combinations of the moment functions of the input process $\xi(t)$. However, the formulas for the moment functions of the output process involve higher-order moment functions of the input process. This is one of the characteristic features of nonlinear transformations as compared with linear transformations. We now consider some examples.

Example 1. Let the nonlinear device have the transfer characteristic

$$y(t) = g(\xi) = a_2\xi^2(t) - a_4\xi^4(t)\,, \tag{8.27}$$

and let the input $\xi(t)$ be a stationary Gaussian process with zero mean and correlation function

$$k(\tau) = \sigma^2 R(\tau)\,.$$

We wish to calculate the mean value and correlation function of $y(t)$. It is easy to see that the first and second-order moment functions of $y(t)$ are

$$\begin{aligned} \langle y(t)\rangle &= a_2\langle\xi^2(t)\rangle - a_4\langle\xi^4(t)\rangle\,, \\ \langle y(t_1)\,y(t_2)\rangle &= a_2^2\langle\xi^2(t_1)\,\xi^2(t_2)\rangle + a_4^2\langle\xi^4(t_1)\,\xi^4(t_2)\rangle \\ &\quad - a_2a_4[\langle\xi^2(t_1)\,\xi^4(t_2)\rangle + \langle\xi^4(t_1)\,\xi^2(t_2)\rangle]\,. \end{aligned} \tag{8.28}$$

As is well known, the univariate moments of a Gaussian process like $\xi(t)$, with mean zero and variance σ^2, are given by the formula

$$\langle\xi^j\rangle = \begin{cases} 1\cdot 3\cdot 5\cdot\cdots\cdot(j-1)\,\sigma^j & \text{for even } j\,, \\ 0 & \text{for odd } j\,. \end{cases} \tag{8.29}$$

To calculate the bivariate moments of the normal process $\xi(t)$, we use the expansion (3.14), from which it follows that

$$\langle\xi^j\xi_\tau^k\rangle = \sigma^{j+k}\sum_{l=0}^{\infty} N_{jl}N_{kl}\,\frac{R^l(\tau)}{l!}\,, \tag{8.30}$$

where

$$N_{jk} = \int_{-\infty}^{\infty} x^j F^{(k+1)}(x)\,dx\,.$$

The set of coefficients N_{jk} forms a matrix of the form

j \ k	0	1	2	3	4	5	6
0	1	0	0	0	0	0	0
1	0	−1	0	0	0	0	0
2	1	0	2·1	0	0	0	0
3	0	−3·1	0	−3·2·1	0	0	0
4	3·1	0	4·3·1	0	4·3·2·1	0	0
5	0	−5·3·1	0	−5·4·3·1	0	−5·4·3·2·1	0
6	5·3·1	0	6·5·3·1	0	6·5·4·3·1	0	6·5·4·3·2·1

This matrix has the following properties, which are easily verified:

a) All elements above the principal diagonal ($j < k$) vanish;

b) The only nonzero elements on or below the principal diagonal ($j \geqslant k$) are those whose indices j and k have the same parity;

c) The nonzero elements N_{jk} are obtained by first forming the product

$$j(j-1)(j-2)\cdots(j-k+1)$$

containing k factors, and then multiplying the result by all odd numbers from $j-k-1$ to 1, e.g.,

$$N_{62} = 6 \cdot 5 \cdot 3 \cdot 1 ;$$

d) The sign of N_{jk} is negative if k is odd and positive if k is even.

Applying these rules and (8.30), we find that

$$\begin{aligned} \langle \xi^2 \xi_\tau^2 \rangle &= \sigma^4[1 + 2R^2(\tau)] , \\ \langle \xi^2 \xi_\tau^4 \rangle &= \sigma^6[3 + 12R^2(\tau)] , \\ \langle \xi^4 \xi_\tau^4 \rangle &= \sigma^8[9 + 72R^2(\tau) + 24R^4(\tau)] . \end{aligned} \tag{8.31}$$

If we use (8.29) and (8.31), then, according to (8.28),

$$\begin{aligned} \langle y \rangle &= a_2\sigma^2 + 3a_4\sigma^4, \\ \langle yy_\tau \rangle &= a_2^2\sigma^4(1 + 2R^2) - 6a_2a_4\sigma^6(1 + 4R^2) + a_4^2\sigma^8(9 + 72R^2 + 24R^4). \end{aligned} \tag{8.32}$$

In particular, the variance of y is

$$Dy = \langle y^2 \rangle - \langle y \rangle^2 = 2a_2^2\sigma^4 - 24a_2a_4\sigma^6 + 96a_4^2\sigma^8. \tag{8.33}$$

Example 2. Let the nonlinear transformation have the parabolic transfer characteristic

$$y(t) = a_1x(t) + a_2x^2(t), \tag{8.34}$$

and let the input signal $x(t)$ be the sum

$$x(t) = s(t) + \xi(t) \tag{8.35}$$

of the "noise" $\xi(t)$ and the "useful signal"

$$s(t) = E\cos(\omega_0 t + \varphi_0), \tag{8.36}$$

which is a sine wave of fixed amplitude and uniformly distributed initial phase. As in the preceding example, we assume that $\xi(t)$ is a stationary Gaussian process, with zero mean and correlation function $k(\tau) = \sigma^2R(\tau)$.

Substituting (8.35) into (8.34), multiplying the result evaluated at time t by the similar expression evaluated at time $t + \tau$, and then taking the average, we obtain

$$\begin{aligned} \langle y \rangle &= a_2\langle s^2 \rangle + a_2\langle \xi^2 \rangle, \\ \langle yy_\tau \rangle &= a_1^2\langle ss_\tau \rangle + a_1^2\langle \xi\xi_\tau \rangle + a_2^2\langle s^2s_\tau^2 \rangle + 4a_2^2\langle ss_\tau \rangle \langle \xi\xi_\tau \rangle \\ &\quad + 2a_2^2\langle s^2 \rangle \langle \xi^2 \rangle + a_2^2\langle \xi^2\xi_\tau^2 \rangle. \end{aligned} \tag{8.37}$$

In deriving the relations (8.37), we have used the fact that the

statistical independence of the random functions $s(t)$ and $\xi(t)$ implies the relations

$$\langle s\xi_\tau\rangle = \langle s\rangle\langle\xi_\tau\rangle, \qquad \langle s^2\xi_\tau^2\rangle = \langle s^2\rangle\langle\xi_\tau^2\rangle,$$

$$\langle ss_\tau\xi\xi_\tau\rangle = \langle ss_\tau\rangle\langle\xi\xi_\tau\rangle, \qquad \langle s\xi_\tau^2\rangle = \langle s\rangle\langle\xi_\tau^2\rangle, \ldots$$

Then, since

$$\langle s\rangle = \langle s_\tau\rangle = 0, \qquad \langle\xi\rangle = \langle\xi_\tau\rangle = 0,$$

we have

$$\langle s\xi_\tau\rangle = 0, \qquad \langle s\xi_\tau^2\rangle = 0, \qquad \langle s^2s_\tau\xi_\tau\rangle = 0, \ldots$$

We have also used the relations

$$\langle s^2s_\tau\rangle = 0, \qquad \langle\xi^2\xi_\tau\rangle = 0,$$

from which it follows that

$$\langle xx_\tau^2\rangle = \langle x_\tau x^2\rangle = 0.$$

Since we already know that

$$\langle\xi\xi_\tau\rangle = \sigma^2R(\tau), \qquad \langle\xi^2\xi_\tau^2\rangle = \sigma^4[1 + 2R^2(\tau)]$$

[cf. the first of the formulas (8.31)], then, to evaluate (8.37) we need only calculate the moments of the useful signal (8.36). Transforming the expression

$$\langle ss_\tau\rangle = E^2\langle\cos(\omega_0t + \varphi_0)\cos(\omega_0t + \omega_0\tau + \varphi_0)\rangle$$

into the form

$$\langle ss_\tau\rangle = \tfrac{1}{2}E^2\langle\cos(2\omega_0t + \omega_0\tau + 2\varphi_0) + \cos\omega_0\tau\rangle,$$

we observe that the first term in the right-hand side vanishes when averaged with respect to the initial phase φ_0, whereas the second term is not random and remains unchanged. Therefore, we have

$$\langle ss_\tau\rangle = \tfrac{1}{2}E^2\cos\omega_0\tau, \tag{8.38}$$

and similarly

$$\langle s^2 s_\tau^2 \rangle = \tfrac{1}{4} E^4 \langle [1 + \cos 2(\omega_0 t + \varphi_0)]\,[1 + \cos 2(\omega_0 t + \omega_0 \tau + \varphi_0)] \rangle$$
$$= \tfrac{1}{4} E^4 (1 + \tfrac{1}{2} \cos 2\omega_0 \tau)\,. \tag{8.39}$$

Substituting (8.38) and (8.39) into (8.37), we obtain

$$\langle y \rangle = \tfrac{1}{2} a_2 E^2 + a_2 \sigma^2\,,$$
$$\langle y y_\tau \rangle = \tfrac{1}{2} a_1^2 E^2 \cos \omega_0 \tau + a_1^2 \sigma^2 R(\tau) + \tfrac{1}{4} a_2^2 E^4 (1 + \tfrac{1}{2} \cos 2\omega_0 \tau) \tag{8.40}$$
$$+ 2a_2^2 \sigma^2 E^2 R(\tau) \cos \omega_0 \tau + a_2^2 \sigma^2 E^2 + a_2^2 \sigma^4 [1 + 2R^2(\tau)]\,,$$

from which it follows at once that the correlation function of the output signal is

$$k_y(\tau) = \tfrac{1}{2} a_1^2 E^2 \cos \omega_0 \tau + a_1^2 \sigma^3 R(\tau) + \tfrac{1}{8} a_2^2 E^4 \cos 2\omega_0 \tau$$
$$+ 2a_2^2 \sigma^3 E^2 R(\tau) \cos \omega_0 \tau + 2a_2^2 \sigma^4 R^2(\tau)\,. \tag{8.41}$$

Example 3. Let the transfer characteristic of the nonlinear device be

$$y = a_1(A - c) + a_2(A - c)^2 \tag{8.42}$$

or

$$y - a_0' = a_1' A + a_2 A^2\,, \tag{8.43}$$

where

$$a_0' = a_2 c^2 - a_1 c\,, \qquad a_1' = a_1 - 2a_2 c\,.$$

Here, we assume that the input $A(t)$ is a Rayleigh process, which is the envelope of some narrow-band Gaussian process (see Chap. 7, Sec. 3), and has the distribution described by (7.59) and (7.60). From (8.43), we find in the usual way that

$$\langle y \rangle = a_0' + a_1' \langle A \rangle + a_2 \langle A^2 \rangle\,, \tag{8.44}$$

$$\langle (y - a_0')(y_\tau - a_0') \rangle = a_1'^2 \langle A A_\tau \rangle + a_1' a_2 [\langle A A_\tau^2 \rangle + \langle A^2 A_\tau \rangle] + a_2^2 \langle A^2 A_\tau^2 \rangle.$$

To calculate the moments of the random function $A(t)$ which appears in (8.44), we use the distributions (7.59) and (7.60). First, we introduce new variables

$$z = \frac{A^2}{2\sigma^2}, \qquad z_\tau = \frac{A_\tau^2}{2\sigma^2},$$

in terms of which

$$A = \sigma(2z)^{1/2}, \qquad A_\tau = \sigma(2z_\tau)^{1/2}. \tag{8.45}$$

Using (7.59) and (7.60), we see that the one and two-dimensional distributions of the random variables z and z_τ are

$$w(z) = e^{-z}, \tag{8.46}$$

$$w(z, z_\tau) = \frac{1}{1-Q^2} I_0\left(\frac{2Q}{1-Q^2}\sqrt{zz_\tau}\right) \exp\left\{-\frac{z+z_\tau}{1-Q^2}\right\}. \tag{8.47}$$

It can be shown[3] that $w(z, z_\tau)$ has the series expansion

$$w(z, z_\tau) = e^{-z-z_\tau} \sum_{n=0}^{\infty} L_n(z)\, L_n(z_\tau) \frac{Q^{2n}}{(n!)^2} \tag{8.48}$$

[cf. formula (4.74)], involving the Laguerre polynomials

$$L_n(z) = e^z \frac{d^n}{dz^n}(z^n e^{-z}), \tag{8.49}$$

which are orthogonal with weight e^{-z} on the interval $(0, \infty)$.

It follows from (8.45) and the formula (8.46) for the one-dimensional probability density that the univariate moments of A are given by the formula

$$\langle A^r \rangle = 2^{r/2}\sigma^r \langle z^{r/2} \rangle = 2^{r/2}\sigma^r \int_0^\infty e^{-z} z^{r/2}\, dz = 2^{r/2}\sigma^r \Gamma\left(\frac{r}{2}+1\right) = 2^{r/2}\sigma^r \left(\frac{r}{2}\right)!, \tag{8.50}$$

[3] The expansion in question is a special case of the *Hille-Hardy formula*. See Bateman Manuscript Project, *Higher Transcendental Functions, Vol. II*, McGraw-Hill Book Co., Inc. (1953), formula (20), p. 189.

where we have used the familiar integral representation of the gamma function (factorial). To calculate the bivariate moments, we use the expansion (8.48) which implies that

$$\langle A^r A^s \rangle = 2^{(r+s)/2}\sigma^{r+s}\langle z^{r/2} z_\tau^{s/2}\rangle = 2^{(r+s)/2}\sigma^{r+s}\sum_{n=0}^{\infty} h_{rn}h_{sn}\frac{Q^{2n}}{(n!)^2}, \qquad (8.51)$$

where h_{rn} denotes the quantity

$$h_{rn} = \int_0^\infty z^{r/2}\, e^{-z}\, L_n(z)\, dz\,. \qquad (8.52)$$

Substituting (8.49) into (8.52) and integrating by parts n times, we find that

$$h_{rn} = (-1)^n \frac{r}{2}\left(\frac{r}{2}-1\right)\cdots\left(\frac{r}{2}-n+1\right)\int_0^\infty e^{-z}\, z^{r/2}\, dz = (-1)^n \frac{[(\frac{1}{2}r)!]^2}{(\frac{1}{2}r-n)!}. \qquad (8.53)$$

If r is even, then the quantities h_{rn} vanish for $n > r/2$. Therefore, the expansion (8.51) contains only a finite number of terms, provided that at least one of the numbers r and s is even. According to (8.53), when $r = 2$ we have

$$h_{20} = 1\,, \qquad h_{21} = -1\,, \qquad h_{22} = h_{23} = \ldots = 0\,. \qquad (8.54)$$

When $r = 1$, all of the coefficients (8.53) are nonzero:

$$h_{10} = \frac{1}{2}! = \frac{\sqrt{\pi}}{2}, \qquad h_{11} = -\frac{1}{2}\frac{\sqrt{\pi}}{2}, \qquad h_{12} = \frac{1}{2}\left(-\frac{1}{2}\right)\frac{\sqrt{\pi}}{2},$$

$$h_{13} = -\frac{1}{2}\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\frac{\sqrt{\pi}}{2}, \ldots, \qquad (8.55)$$

$$|\,h_{1n}\,| = (2n-3)(2n-5)\cdots 3\cdot 1\,\frac{\sqrt{\pi}}{2^{n+1}}, \ldots$$

Substituting (8.54) and (8.55) into (8.51), we obtain

$$\begin{aligned}
\langle A^2 A_\tau^2 \rangle &= 4\sigma^4(1 + Q^2), \\
\langle A A_\tau^2 \rangle = \langle A^2 A_\tau \rangle &= \sigma^3 \sqrt{2\pi}\left(1 + \frac{1}{2}Q^2\right), \\
\langle A A_\tau \rangle &= \frac{\pi}{2}\sigma^3\left(1 + \frac{1}{4}Q^2 + \frac{1}{64}Q^4 + \frac{1}{256}Q^6 + \ldots\right)
\end{aligned} \tag{8.56}$$

Then, using (8.44), we arrive at the final result:

$$\begin{aligned}
\langle y \rangle &= a_0' + a_1'\sigma\sqrt{\frac{\pi}{2}} + 2a_2\sigma^2, \\
\langle (y - a_0)(y_\tau - a_0) \rangle &= \frac{\pi}{2}a_1'^2\sigma^2\left(1 + \frac{1}{4}Q^2 + \frac{1}{64}Q^4 + \frac{1}{256}Q^6 + \ldots\right) \\
&\quad + 2a_1'a_2\sigma^3\sqrt{2\pi}\left(1 + \frac{1}{2}Q^2\right) + 4a_2^2\sigma^4(1 + Q^2), \\
k_y(\tau) &= \frac{\pi}{8}a_1'^2\sigma^2\left(Q^2 + \frac{1}{16}Q^4 + \frac{1}{64}Q^6 + \ldots\right) \\
&\quad + \sqrt{2\pi}\, a_1'a_2\sigma^3 Q^2 + 4a_2^2\sigma^4 Q^2.
\end{aligned} \tag{8.57}$$

3. Moment Functions for Piecewise Linear Transfer Characteristics. The Direct Method

The moment functions of random signals subjected to piecewise linear transformations can be calculated by two intimately related methods, one involving the use of series expansions of the two-dimensional probability density (the *direct method*) and the other involving the characteristic function and the evaluation of integrals (*Rice's method*). In the next two sections, we shall use typical examples to illustrate the gist of these two methods, in keeping with our policy of deriving quantitative results of practical interest.

3.1. Gaussian input processes. First, we calculate the mean value $\langle \eta \rangle$ and the correlation function $k_\eta(\tau)$ of the process $\eta(t)$ obtained at the output of a nonlinear device with a piecewise linear transfer

characteristic $g(\xi)$, of the kind shown in Figure 6, when the input is a stationary Gaussian process $\xi(t)$ with zero mean and correlation function

$$k(\tau) = \sigma^2 R(\tau).$$

In this case, the one-dimensional probability density is

$$w(\xi) = \frac{1}{\sigma} F'\left(\frac{\xi}{\sigma}\right), \tag{8.58}$$

where

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dy,$$

and the two-dimensional probability density $w(\xi, \xi_\tau)$ can be conveniently written in the form (3.14).

Thus, we assume that the function $g(\xi)$ is continuous and piecewise linear, and that its derivative $g'(\xi)$ is discontinuous at the points $c_1, c_2, \ldots, c_r$ (see Figure 6). This means that

$$g''(\xi) = \sum_{i=1}^{r} b_i \delta(\xi - c_i). \tag{8.59}$$

Then, the mean value

$$\langle \eta \rangle = \int g(\xi)\, w(\xi)\, d\xi = \frac{1}{\sigma} \int g(\xi)\, F'\left(\frac{\xi}{\sigma}\right) d\xi \tag{8.60}$$

can be written in the form

$$\langle \eta \rangle = \lim_{\xi_1 \to \infty} \int_{-\infty}^{\xi_1} g(\xi)\, dF\left(\frac{\xi}{\sigma}\right). \tag{8.61}$$

Integrating by parts, we obtain

$$\int_{-\infty}^{\xi_1} g(\xi)\, dF\left(\frac{\xi}{\sigma}\right) = g(\xi_1) F\left(\frac{\xi_1}{\sigma}\right) - \int_{-\infty}^{\xi_1} g'(\xi) F\left(\frac{\xi}{\sigma}\right) d\xi. \tag{8.62}$$

We now introduce the function

$$F^{(-1)}(x) = \int_{-\infty}^{x} F(x)\, dx = xF(x) + F'(x), \tag{8.63}$$

which has the properties

$$\frac{d}{dx}F^{(-1)}(x) = F(x), \qquad F^{(-1)}(-\infty) = 0.$$

Using these properties, and carrying out a second integration by parts in (8.62), we obtain

$$\int_{-\infty}^{\xi_1} g(\xi)\, dF\left(\frac{\xi}{\sigma}\right) = g(\xi_1) F\left(\frac{\xi_1}{\sigma}\right) - \sigma \int_{-\infty}^{\xi_1} g'(\xi)\, dF^{(-1)}\left(\frac{\xi}{\sigma}\right) \tag{8.64}$$

$$= g(\xi_1) F\left(\frac{\xi_1}{\sigma}\right) - \sigma g'(\xi_1) F^{(-1)}\left(\frac{\xi_1}{\sigma}\right) + \sigma \int_{-\infty}^{\xi_1} g''(\xi) F^{(-1)}\left(\frac{\xi}{\sigma}\right) d\xi.$$

According to (8.61), the mean value $\langle \eta \rangle$ can be obtained by taking the limit of (8.64) as $\xi_1 \to \infty$. Since, as $\xi_1 \to \infty$, $F(\xi_1/\sigma)$ rapidly approaches 1, while $F^{(-1)}(\xi_1/\sigma)$ approaches ξ_1/σ [cf. (8.63)], it follows that

$$\langle \eta \rangle = \lim_{\xi_1 \to \infty} [g(\xi_1) - \xi_1 g'(\xi_1)] + \sigma \int_{-\infty}^{\infty} g''(\xi) F^{(-1)}\left(\frac{\xi}{\sigma}\right) d\xi,$$

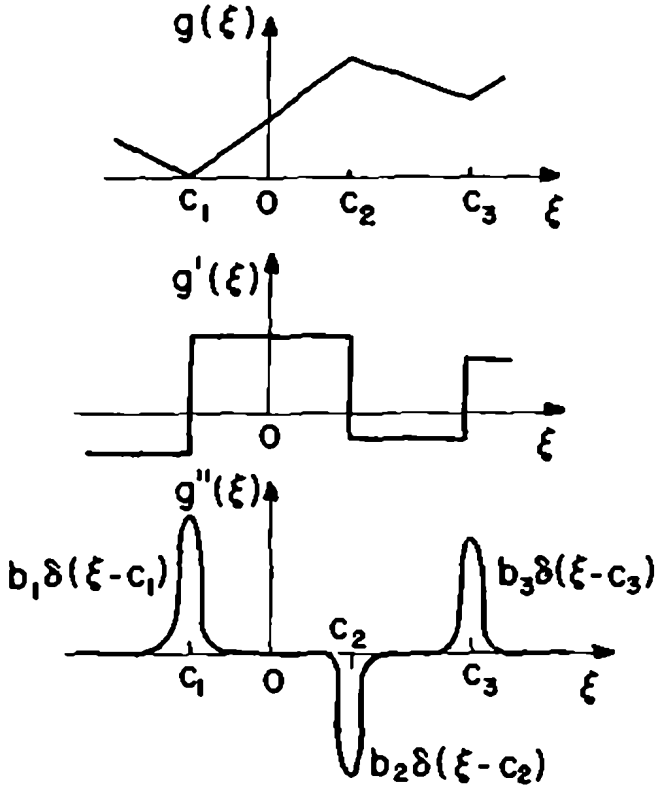

Fig. 6. A piecewise linear transfer characteristic and its derivatives.

or

$$\langle \eta \rangle = \lim_{\xi_1 \to \infty} [g(\xi) - \xi g'(\xi)] + \sigma \sum_{l=1}^{r} b_l F^{(-1)}\left(\frac{c_l}{\sigma}\right), \tag{8.65}$$

where we have used (8.59). Moreover, from (3.14) we find that the second moment of $\eta(t)$ can be written as a series of the form

$$\langle \eta \eta_\tau \rangle = \sum_{n=0}^{\infty} \left[\frac{1}{\sigma} \int_{-\infty}^{\infty} g(\xi) F^{(n+1)}\left(\frac{\xi}{\sigma}\right) d\xi\right]^2 \frac{R^n(\tau)}{n!},$$

whose first term is just $\langle \eta \rangle^2$, according to (8.60). Subtracting this first term from $\langle \eta \eta_\tau \rangle$, we obtain the correlation function

$$k_\eta(\tau) = \sum_{n=1}^{\infty} \left[\int_{-\infty}^{\infty} g(\xi) F^{(n+1)}\left(\frac{\xi}{\sigma}\right) \frac{d\xi}{\sigma}\right]^2 \frac{R^n(\tau)}{n!}. \tag{8.66}$$

To evaluate each of the integrals in (8.66), we again integrate by parts twice. For $n = 1$, we obtain first

$$\int_{-\infty}^{\infty} g(\xi) F''\left(\frac{\xi}{\sigma}\right) \frac{d\xi}{\sigma} = -\int_{-\infty}^{\infty} g'(\xi) F'\left(\frac{\xi}{\sigma}\right) d\xi,$$

and then

$$\int_{-\infty}^{\infty} g(\xi) F''\left(\frac{\xi}{\sigma}\right) \frac{d\xi}{\sigma} = -\sigma g'(\infty) F(\infty) + \sigma \int_{-\infty}^{\infty} g''(\xi) F\left(\frac{\xi}{\sigma}\right) d\xi. \tag{8.67}$$

For $n > 1$, the integrals can be evaluated even more simply, since the corresponding integrated terms vanish at $\pm\infty$. Finally, substituting (8.59) into (8.67), we find that

$$\int_{-\infty}^{\infty} g(\xi) F''\left(\frac{\xi}{\sigma}\right) \frac{d\xi}{\sigma} = -\sigma g'(\infty) + \sigma \sum_{l=1}^{r} b_l F\left(\frac{c_l}{\sigma}\right),$$

$$\int_{-\infty}^{\infty} g(\xi) F^{(n+1)}\left(\frac{\xi}{\sigma}\right) \frac{d\xi}{\sigma} = \sigma \sum_{l=1}^{r} b_l F^{(n-1)}\left(\frac{c_l}{\sigma}\right) \qquad (n > 1).$$

We must now substitute the expressions just obtained into formula (8.66). A variety of special cases are possible, as shown in Figure 7. If $g(\xi)$ consists of a single line segment, so that $r = 0$, $b_l = 0$, then only the first term remains in (8.66). This case

corresponds to amplification of the signal $|g'|$ times, together with a possible change in the d-c component of the signal. Next, suppose that $g(\xi)$ consists of two line segments, joined at the point $\xi = c_1$, where $g' = \alpha$ to the left of c_1 and $g' = \beta$ to the right of c_1 (see Figure 7a). Then

$$g'(\infty) = \beta\,, \qquad b_1 = \beta - \alpha\,,$$

and we have

$$k_\eta(\tau) = \sigma^2\left[\beta + (\alpha - \beta)F\left(\frac{c_1}{\sigma}\right)\right]^2 R(\tau) + (\beta - \alpha)^2 \sum_{n=2}^{\infty}\left[F^{(n-1)}\left(\frac{c_1}{\sigma}\right)\right]^2 \frac{R^n(\tau)}{n!}. \tag{8.68}$$

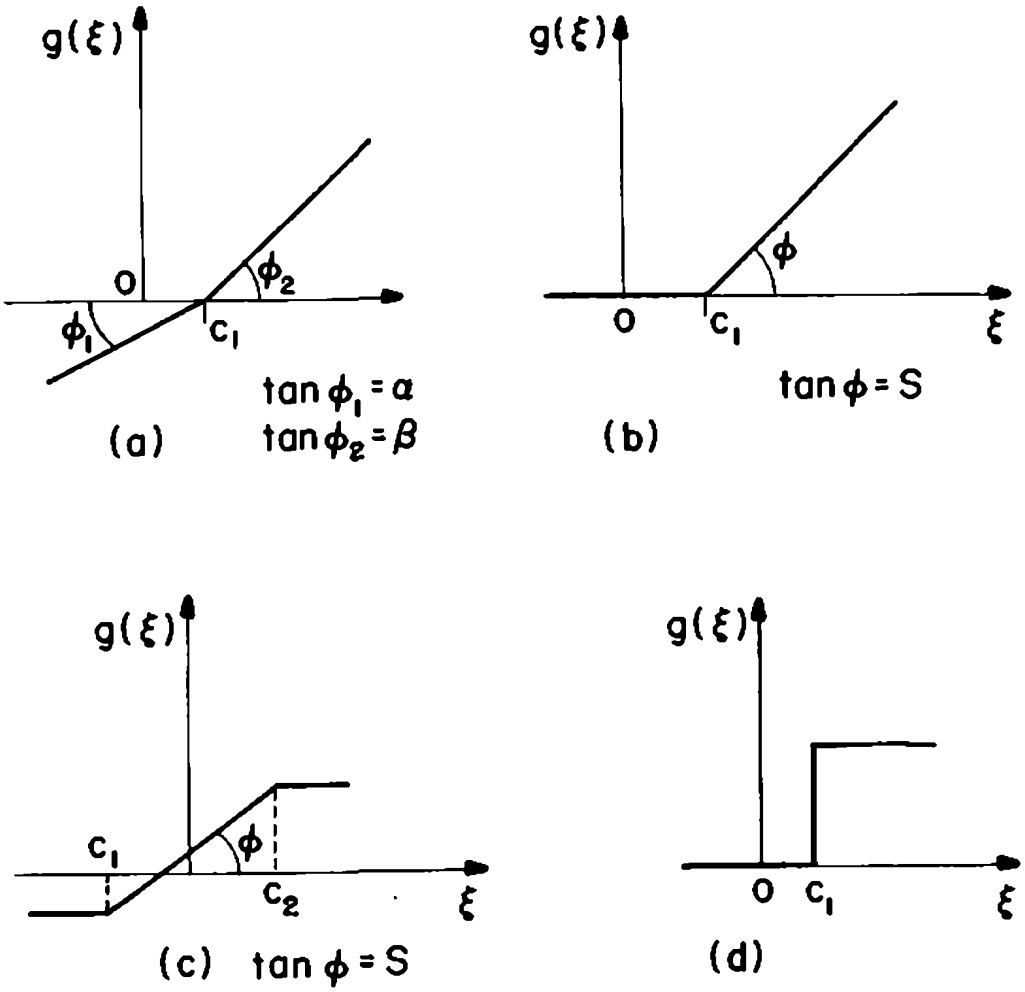

Fig. 7. Typical piecewise linear transfer characteristics.

In particular, if the "break point" is centrally located, i.e., $c_1 = 0$, then

$$k_\eta(\tau) = \frac{\sigma^2}{4}(\alpha+\beta)^2 R(\tau) + \sigma^2(\beta-a)^2 \sum_{n=2}^{\infty} \left[F^{(n-1)}(0)\right]^2 \frac{R^n(\tau)}{n!} \tag{8.69}$$

$$= \frac{\sigma^2}{4}(\alpha+\beta)^2 R(\tau) + \frac{\sigma^2}{2\pi}(\beta-\alpha)^2 (R \text{ arc sin } R + \sqrt{1-R^2} - 1).$$

By examining the case

$$g(\xi) = \begin{cases} S(\xi - c_1) & \text{for } \xi > c_1, \\ 0 & \text{for } \xi < c_1, \end{cases} \tag{8.70}$$

(see Figure 7b) where

$$\alpha = 0, \qquad \beta = S,$$

we can consider "peaks" of the random function $\xi(t)$ and we can calculate the degree of correlation between them. In this case, we have

$$k_\eta(\tau) = S^2\sigma^2\left\{F^2\left(-\frac{c_1}{\sigma}\right) R(\tau) + \sum_{n=2}^{\infty} \frac{1}{n!}\left[F^{(n-1)}\left(\frac{c_1}{\sigma}\right)\right]^2 R^n(\tau)\right. . \tag{8.71}$$

In Figure 8, we show the results of calculations of the correlation coefficient

$$R_\eta(\tau) = \frac{k_\eta(\tau)}{k_\eta(0)}$$

based on this formula, when terms up to $n = 7$ are retained in the series and when the correlation function of the input process is $R(\tau) = e^{-\alpha\tau^2}$ (shown as the dashed curve). Curves corresponding to five values of the threshold $\gamma = c_1/\sigma$ are plotted.

As an example of a transfer characteristic with two break points, we consider the *asymmetric limiter*

$$\eta = g(\xi) = \begin{cases} Sc_2 & \text{for } \xi > c_2, \\ S\xi & \text{for } c_1 < \xi < c_2, \\ Sc_1 & \text{for } \xi < c_1 \end{cases} \tag{8.72}$$

(see Figure 7c), where

$$g'(\infty) = 0, \qquad b_1 = S, \qquad b_2 = -S.$$

In this case, the correlation function is

$$k_\eta(\tau) = S^2\sigma^2 \sum_{n=1}^{\infty} \left[F^{(n-1)}\left(\frac{c_2}{\sigma}\right) - F^{(n-1)}\left(\frac{c_1}{\sigma}\right)\right]^2 \frac{R^n(\tau)}{n!}, \qquad (8.73)$$

and, according to (8.65) and (8.63), the mean value is

$$\langle\eta\rangle = Sc_2 + S\left[c_1F\left(\frac{c_1}{\sigma}\right) - c_2F\left(\frac{c_2}{\sigma}\right) + \sigma F'\left(\frac{c_1}{\sigma}\right) - \sigma F'\left(\frac{c_2}{\sigma}\right)\right] \qquad (8.74)$$

To find the variance of the output signal in each of the examples just discussed, we can set $\tau = 0$, $R = 1$ in formulas (8.68), (8.69), (8.71) and (8.73); this gives the variance in the form of an infinite series. However, we note that it is simpler to find the variance in the ordinary way, by using the one-dimensional probability density $w(\xi)$ to calculate $\langle\eta^2\rangle - \langle\eta\rangle^2$, since this will give the result in finite terms.

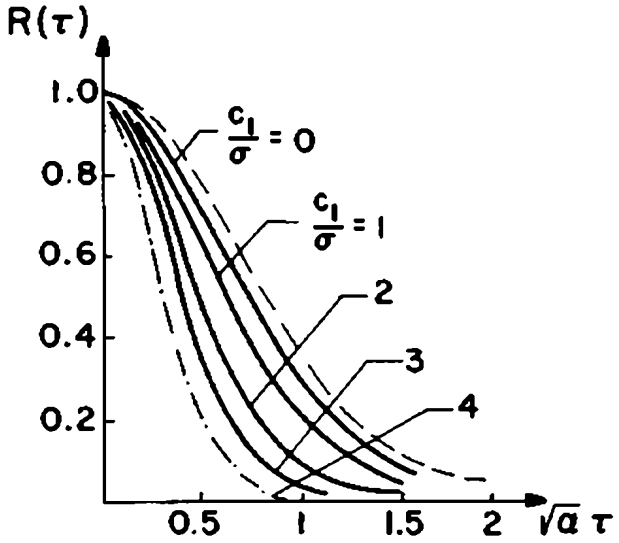

Fig. 8. The correlation coefficient $R_\eta(\tau)$ corresponding to (8.71), when $R(\tau) = e^{-\alpha\tau^2}$.

The method used above, which is based on integration by parts, is also effective for calculating higher moments. The method can also be generalized to the case of piecewise continuous transfer characteristics of more general form, e.g., characteristics with

jump discontinuities or characteristics made up of parabolic pieces. For example, if the transfer characteristic is

$$\eta = \begin{cases} q & \text{for} \quad \xi > c_1, \\ 0 & \text{for} \quad \xi < c_1 \end{cases} \tag{8.75a}$$

(see Figure 7d), then a single integration by parts is enough, while if

$$\eta = \begin{cases} \beta(\xi - c_1)^2 & \text{for} \quad \xi > c_1, \\ 0 & \text{for} \quad \xi < c_1, \end{cases} \tag{8.75b}$$

then three integrations by parts are needed. The resulting expressions for the correlation function are

$$k_\eta(\tau) = q^2 \sum_{n=1}^{\infty} \frac{1}{n!} \left[F^{(n)}\left(\frac{c_1}{\sigma}\right)\right]^2 R^n(\tau) \tag{8.76a}$$

and

$$k_\eta(\tau) = 4\beta^2\sigma^4 \sum_{n=1}^{\infty} \frac{1}{n!} \left[F^{(n-2)}\left(-\frac{c_1}{\sigma}\right)\right]^2 R^n(\tau), \tag{8.76b}$$

respectively.

3.2. Rayleigh input processes. We now consider two cases where the input signal is a Rayleigh process and the transfer characteristic is piecewise linear.

Example 1. Suppose the input process [whose one-dimensional probability density is the Rayleigh distribution (7.59)] undergoes the transformation

$$\eta = \begin{cases} q & \text{for} \quad A > A_0, \\ 0 & \text{for} \quad A < A_0. \end{cases} \tag{8.77}$$

Introducing the process $z(t) = A^2(t)/2\sigma^2$, as before, we have

$$\eta = \begin{cases} q & \text{for} \quad z > z_0, \\ 0 & \text{for} \quad z < z_0, \end{cases} \tag{8.78}$$

where

$$z_0 = \frac{A_0^2}{2\sigma^2}.$$

Then, using the distributions (8.47) and (8.48) of the process $z(t)$, we find that

$$\langle \eta \rangle = q \int_{z_0}^{\infty} e^{-z}\, dz = q e^{-z_0}, \tag{8.79}$$

$$\langle \eta \eta_\tau \rangle = q^2 \sum_{n=0}^{\infty} \frac{Q^{2n}}{(n!)^2} \left[\int_{z_0}^{\infty} L_n(z)\, e^{-z}\, dz \right]^2. \tag{8.80}$$

To evaluate the integrals appearing in (8.80), we use the definition of the Laguerre polynomials, obtaining

$$\int_{z_0}^{\infty} L_n(z)\, e^{-z}\, dz = -\frac{d^{n-1}}{dz^{n-1}} (z^n e^{-z}) \Big|_{z=z_0},$$

which can be written as

$$\int_{z_0}^{\infty} L_n(z)\, e^{-z}\, dz = -z_0 L_{n-1}^{(1)}(z_0)\, e^{-z_0}, \tag{8.81}$$

in terms of the *associated Laguerre polynomials*

$$L_n^{(1)}(z) = z^{-1} e^z \frac{d^n}{dz^n} (z^{n+1} e^{-z}). \tag{8.82}$$

Therefore, according to (8.79) and (8.80), the output process, which is actually a sequence of pulses, has the correlation function

$$k_\eta(\tau) = q^2 e^{-2z_0} z_0^2 \sum_{n=1}^{\infty} \frac{Q^{2n}(\tau)}{(n!)^2} [L_{n-1}^{(1)}(z_0)]^2. \tag{8.83}$$

For reference, we write down the first few of the polynomials (8.82):

$$L_0^{(1)}(z) = 1, \quad L_1^{(1)}(z) = 2 - z, \quad L_2^{(1)}(z) = 6 - 6z + z^2,$$
$$L_3^{(1)}(z) = 24 - 36z + 12z^2 - z^3, \ldots \tag{8.84}$$

Using (8.84) and substituting $z_0 = A_0^2/2\sigma^2$ into (8.83), we obtain

$$k_\eta(\tau) = q^2 \frac{A_0^4}{4\sigma^4} e^{-A_0^2/\sigma^2} \left[Q^2 + \left(1 - \frac{A_0^2}{4\sigma^2}\right)^2 Q^4 + \left(1 - \frac{A_0^2}{2\sigma^2} + \frac{A_0^4}{24\sigma^4}\right)^2 Q^6 + \ldots \right]. \tag{8.85}$$

We could calculate the variance of $\eta(t)$ by setting $\tau = 0, Q = 1$ in (8.85). However, it is simpler to make direct use of the one-dimensional probability density; this leads immediately to the formulas

$$\begin{aligned} \langle \eta^2 \rangle &= q^2 e^{-z_0}, \\ \mathbf{D}\eta &= k_\eta(0) = q^2 e^{-2z_0}(e^{z_0} - 1). \end{aligned} \tag{8.86}$$

Dividing (8.83) or (8.85) by the variance $\mathbf{D}\eta$, we find that the correlation coefficient is

$$R_\eta(\tau) = \frac{k_\eta(\tau)}{k_\eta(0)} = B_0[Q^2(\tau) + B_2 Q^4(\tau) + B_3 Q^6(\tau) + \ldots], \tag{8.87}$$

where

$$B_0 = \frac{A_0^4}{4\sigma^2}(e^{A_0^2/2\sigma^2} - 1)^{-1},$$

$$B_2 = \left(1 - \frac{A_0^2}{4\sigma^2}\right)^2,$$

$$B_3 = \left(1 - \frac{A_0^2}{2\sigma^2} + \frac{A_0^4}{24\sigma^4}\right)^2.$$

In Figure 9, we show the results of calculations based on formula

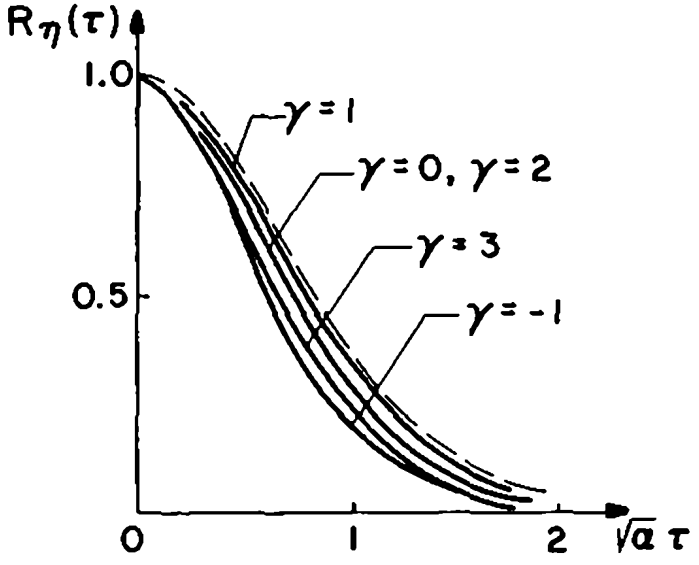

Fig. 9. The correlation coefficient $R_\eta(\tau)$ given by (8.87), when $Q(\tau) = e^{-\alpha\tau^2}$.

(8.87), where only the indicated terms of the series have been retained. These calculations were made for the case where

$$Q(\tau) = e^{-\alpha\tau^2}$$

(shown as the dashed curve). The quantity

$$\gamma = \frac{A_0}{\sigma} - \frac{\langle A \rangle}{\sigma}$$

specifies the operating level of the limiter, and we have plotted the curves corresponding to the values $\gamma = -1, 0, 1, 2$ and 3. It should be noted that

$$\frac{\langle A \rangle}{\sigma} = \sqrt{\frac{\pi}{2}} \approx 1.25$$

for a Rayleigh process.

Example 2. Suppose the input Rayleigh process $A(t)$ undergoes the transformation

$$\eta = \begin{cases} S(A - A_0) & \text{for } A > A_0, \\ 0 & \text{for } A < A_0, \end{cases} \tag{8.88}$$

which "cuts off the envelope peaks" exceeding the threshold $A = A_0$. Using the univariate Rayleigh distribution (7.59), we obtain the mean values

$$\begin{aligned} \langle \eta \rangle &= S\sigma\sqrt{2\pi}\,F(-\alpha), \\ \langle \eta^2 \rangle &= 2S^2\sigma^2[e^{-\alpha^2/2} - \sqrt{2\pi}\,\alpha F(-\alpha)], \end{aligned} \tag{8.89}$$

and the variance

$$\mathbf{D}\eta = S^2\sigma^2[2e^{-\alpha^2/2} - 2\sqrt{2\pi}\,\alpha F(-\alpha) - 2\pi F^2(-\alpha)], \tag{8.90}$$

where $\alpha = A_0/\sigma$.

To find the correlation function $k_\eta(\tau)$, we use the expansion (8.48) of the bivariate distribution. First, we note that the transformation (8.88) is equivalent to the transformation

$$\eta = \sqrt{2}S\sigma \begin{cases} z^{1/2} - z_0^{1/2} & \text{for } z > z_0, \\ 0 & \text{for } z < z_0, \end{cases} \tag{8.91}$$

where, as usual,

$$z = \frac{A^2}{2\sigma^2}, \qquad z_0 = \frac{A_0^2}{2\sigma^2}.$$

After subtracting $\langle \eta \rangle^2$ from (8.48), we have

$$k_\eta(\tau) = 2S^2\sigma^2 \sum_{n=1}^{\infty} \frac{h_n^2}{(n!)^2} Q^{2n}(\tau), \tag{8.92}$$

where

$$h_\eta = \int_{z_0}^{\infty} (z^{1/2} - z_0^{1/2}) L_n(z)\, e^{-z}\, dz.$$

Next, we integrate this last expression by parts, obtaining

$$h_n = -\int_{z_0}^{\infty} z L_{n-1}^{(1)}(z)\, e^{-z}\, dz^{1/2}, \tag{8.93}$$

where we have used the relation

$$L_n(z)\, e^{-z}\, dz = d[z L_{n-1}^{(1)}(z)\, e^{-z}],$$

which is equivalent to (8.81).

Introducing the parameter μ, and then differentiating with respect to μ, we can write (8.93) in the form

$$h_n = \frac{\partial}{\partial \mu} L_{n-1}^{(1)}\left(-\frac{\partial}{\partial \mu}\right) \int_{z_0}^{\infty} e^{-\mu z}\, dz^{1/2} \Big|_{\mu=1}. \tag{8.94}$$

Then, setting $z = A^2/2\sigma^2 = a^2/2$, we express the integral in (8.94) in terms of the probability integral (see p. 42):

$$\int_{z_0}^{\infty} e^{-\mu z}\, dz^{1/2} = \frac{1}{\sqrt{2}} \int_{\alpha}^{\infty} e^{-\mu a^2/2}\, da = \sqrt{\frac{\pi}{\mu}} F(-\alpha\sqrt{\mu}).$$

It follows that

$$h_n = \sqrt{\pi}\, \frac{\partial}{\partial \mu} L_{n-1}^{(1)}\left(-\frac{\partial}{\partial \mu}\right) \left[\frac{1}{\sqrt{\mu}} F(-\alpha\sqrt{\mu})\right] \Big|_{\mu=1}. \tag{8.95}$$

For computational purposes, it is more convenient to use the somewhat different formula

$$h_n = \sqrt{\pi}\,\frac{\partial^n}{\partial \mu^n}\left[\mu^{n-(3/2)}\,F(-\alpha\sqrt{\mu})\right]\Big|_{\mu=1}, \tag{8.96}$$

which can be verified by using the relations[4]

$$-zL_{n-1}^{(1)}(z) = \sum_{l=0}^{n}\binom{n}{l}(n-1)(n-2)\ldots l(-z)^l,$$

$$\sum_{l=0}^{n}\binom{n}{l}(n-1)(n-2)\ldots l\mu^{l-1}\frac{\partial^l F_1}{\partial \mu^l} = \sum_{l=0}^{n}\binom{n}{l}\frac{\partial^{n-l}\mu^{n-1}}{\partial \mu^{n-l}}\frac{\partial^l F_1}{\partial \mu^l}$$

$$= \frac{\partial^n}{\partial \mu^n}(\mu^{n-1}F_1),$$

where

$$F_1 = \frac{1}{\sqrt{\mu}}F(-\alpha\sqrt{\mu}).$$

The first few coefficients h_n are

$$h_1 = -\frac{\sqrt{\pi}}{2}[F(-\alpha) + \alpha F'(\alpha)],$$

$$h_2 = -\frac{\sqrt{\pi}}{4}[F(-\alpha) + \alpha F'(\alpha) + \alpha^2 F''(\alpha)], \tag{8.97}$$

$$h_3 = -\frac{3\sqrt{\pi}}{8}[F(-\alpha) + \alpha F'(\alpha) + 2\alpha^2 F''(\alpha) + \tfrac{1}{3}\alpha^3 F'''(\alpha)].$$

If we retain only the coefficients (8.97) in our calculations, setting all the others equal to zero, then the correlation coefficient of the process $\eta(t)$ [which is just the random sequence of peaks of the envelope $A(t)$ exceeding the threshold A_0] is given by the expression

$$R_\eta(\tau) = B^{-1}\left(h_1^2Q^2 + \frac{1}{4}h_2^2Q^4 + \frac{1}{36}h_3^2Q^6\right), \tag{8.98}$$

[4] Cf. I. M. Ryshik and I. S. Gradstein, *op. cit.*, formula (7.140.1), p. 381.

where

$$B = e^{-\alpha^2/2} - \sqrt{2\pi}\,\alpha F(-\alpha) - \pi F^2(-\alpha).$$

[Cf. formula (8.90).] In Figure 10, we show how $R_\eta(\tau)$ depends on

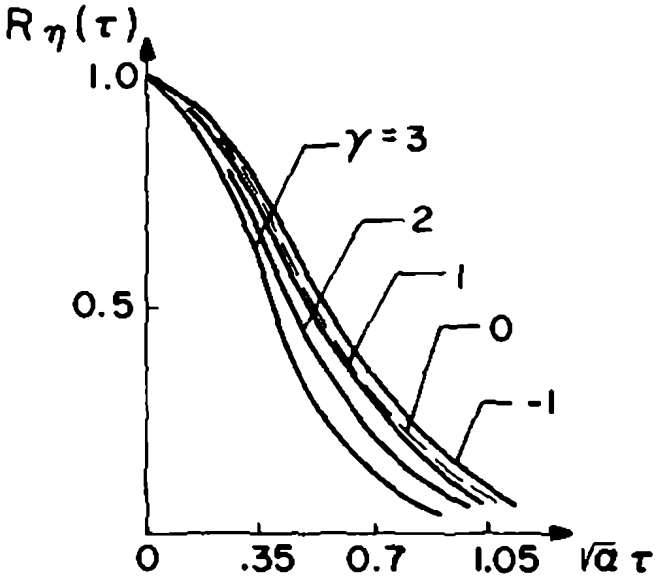

Fig. 10. The correlation coefficient $R_\eta(\tau)$ given by (8.98), when $Q(\tau) = e^{-\alpha\tau^2}$.

the value of the threshold

$$\gamma = \frac{A_0}{\sigma} - 1.25 = -1, 0, 1, 2, 3,$$

in the case where

$$Q(\tau) = e^{-\alpha\tau^2}.$$

It is clear from the figure that the correlation coefficient of the peaks can differ appreciably from the function $Q(\tau)$, shown as the dashed curve in Figure 10.[5] In fact, $R(\tau)$ "spreads out" for small thresholds, and then "shrinks," starting from some value of $\gamma \approx 0$.

[5] $Q(\tau)$ can be regarded as a kind of envelope correlation coefficient, but, as shown by (8.56), $Q(\tau) \neq \langle AA_\tau\rangle - \langle A\rangle^2$.

4. Moment Functions for Exponential Transfer Characteristics. Rice's Method

4.1. The moment functions of the output

$$\eta(t) = he^{a\xi(t)} \tag{8.99}$$

of a nonlinear device with an exponential transfer characteristic can be expressed in terms of the characteristic functions of the input signal $\xi(t)$. This fact can be exploited in cases where the characteristic functions are known, e.g., when $\xi(t)$ is a Gaussian process. By averaging (8.99), we obtain

$$\begin{aligned} \langle \eta \rangle &= h\Theta_1(-ia)\,, \\ \langle \eta\eta_\tau \rangle &= h^2\Theta_2(-ia, -ia)\,, \end{aligned} \tag{8.100}$$

where

$$\begin{aligned} \Theta_1(u) &= \langle e^{iu\xi} \rangle\,, \\ \Theta_2(u_1, u_2) &= \langle e^{iu_1\xi + iu_2\xi_\tau} \rangle \end{aligned} \tag{8.101}$$

are the one and two-dimensional characteristic functions. If $\xi(t)$ is a stationary Gaussian process with zero mean value, then as we know [cf. (3.1)]

$$\begin{aligned} \Theta_1(u) &= \exp\{-\tfrac{1}{2}\sigma^2u^2\}\,, \\ \Theta_2(u_1, u_2) &= \exp\{-\tfrac{1}{2}\sigma^2[u_1^2 + 2R(\tau)\,u_1u_2 + u_2^2]\}\,. \end{aligned} \tag{8.102}$$

Then, we can use (8.100) to find expressions for $\langle \eta \rangle$, $\langle \eta\eta_\tau \rangle$ and the output correlation function

$$k_\eta(\tau) = h^2e^{a^2\sigma^2}[e^{a^2\sigma^2R(\tau)} - 1]\,. \tag{8.103}$$

When the transfer characteristic of the nonlinear device is not exponential, we can still use characteristic functions to calculate the moments of the output process. To do so, we use *Rice's method*, the basic idea of which is to represent the transfer characteristic of the nonlinear device by a contour integral

$$\eta = g(\xi) = \frac{1}{2\pi}\int_L F(i\Omega)\,e^{i\Omega\xi}\,d\Omega\,, \tag{8.104}$$

where L is a suitably chosen contour of integration in the complex Ω-plane. If the function $g(\xi)$ vanishes at infinity ($\xi = \pm\infty$), we can choose (8.104) to be the Fourier transform; then

$$F(i\Omega) = \int_{-\infty}^{\infty} g(\xi)\, e^{-i\xi\Omega}\, d\xi\,, \tag{8.105}$$

and we need only consider real values of Ω. If $g(\Omega)$ vanishes for $\xi < 0$ (or at least for ξ less than some fixed number), we can apply the theory of the Laplace transform, by regarding (8.104) as the inversion integral

$$g(\xi) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} e^{p\xi}\, F(p)\, dp$$

of the transform

$$F(p) = \int_0^{\infty} e^{-p\xi}\, g(\xi)\, d\xi \qquad (p = i\Omega)\,.$$

In this case, if $g(\xi)$ grows no faster than some power of ξ, as $\xi \to +\infty$, we can choose the contour of integration in (8.104) to be the real axis ($c = 0$), with a downward indentation at the origin. (The indentation is to the right in the complex plane corresponding to $p = i\Omega$.) If $g(\xi)$ goes to infinity in both directions (i.e., as $\xi \to \pm\infty$), then it is appropriate to use the theory of the two-sided (bilateral) Laplace transform.

Averaging first (8.104), and then the product of two expressions of the form (8.104) evaluated at the times t and $t + \tau$, we can express the moments $\langle\eta\rangle$ and $\langle\eta\eta_\tau\rangle$ in terms of the characteristic functions of the input process, which are assumed to be known:

$$\begin{aligned}\langle\eta\rangle &= \frac{1}{2\pi}\int_L F(i\Omega)\,\langle\exp\{i\Omega\xi\}\rangle\, d\Omega\,,\\ \langle\eta\eta_\tau\rangle &= \frac{1}{4\pi^2}\int_L\int_L F(i\Omega)\,F(i\Omega_\tau)\,\langle\exp\{i(\Omega\xi + \Omega_\tau\xi_\tau)\}\rangle\, d\Omega\, d\Omega_\tau\,.\end{aligned} \tag{8.106}$$

For example, if $\xi(t)$ is a stationary Gaussian process, then, according to (8.102), we have

$$\langle\eta\eta_\tau\rangle = \frac{1}{4\pi^2}\int_L\int_L F(i\Omega)\,F(i\Omega_\tau)\exp\{-\tfrac{1}{2}\sigma^2(\Omega^2 + 2R\Omega\Omega_\tau + \Omega_\tau^2)\}\, d\Omega\, d\Omega_\tau\,. \tag{8.107}$$

If this integral cannot be evaluated explicitly, we substitute the expansion

$$e^{-\sigma^2 R\Omega\Omega_\tau} = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \sigma^{2n} R^n \Omega^n \Omega_\tau^n \tag{8.108}$$

into (8.107), obtaining

$$\langle \eta\eta_\tau \rangle = \sum_{n=0}^{\infty} \frac{h_n^2}{n!} \sigma^{2n} R^n(\tau), \tag{8.109}$$

where the coefficients

$$h_n = \frac{i^n}{2\pi} \int_L F(i\Omega)\, e^{-\sigma^2\Omega^2/2}\, \Omega^n\, d\Omega \tag{8.110}$$

are much easier to compute. Then, using the first of the formulas (8.106), we subtract $\langle \eta \rangle^2$ from (8.110), which gives the correlation function

$$k_\eta(\tau) = \sum_{n=1}^{\infty} \frac{h_n^2}{n!} \sigma^{2n} R^n(\tau). \tag{8.111}$$

Example 1. To illustrate Rice's method, we find the correlation function $k_\eta(\tau)$ corresponding to the nonlinear transformation

$$\eta(t) = \begin{cases} A\sqrt{1-\xi^2(t)} & \text{for } |\xi| < 1, \\ 0 & \text{for } |\xi| > 1. \end{cases} \tag{8.112}$$

In this case, we can simply take the Fourier transform (8.105), obtaining

$$F(i\Omega) = A \int_{-1}^{1} e^{-i\Omega\xi} \sqrt{1-\xi^2}\, d\xi,$$

where the integral is readily expressed in terms of a Bessel function,[6] i.e,

$$F(i\Omega) = \pi A \frac{J_1(\Omega)}{\Omega}, \tag{8.113}$$

[6] I. M. Ryshik and I. S. Gradstein, *op. cit.*, formula (6.413.3), p. 312.

and the inverse transform is just

$$\eta = \frac{A}{2}\int_{-\infty}^{\infty} \frac{J_1(\Omega)}{\Omega} e^{i\Omega\xi}\, d\Omega\,. \tag{8.114}$$

For Gaussian noise with the characteristic function (8.102), the desired correlation function is given by (8.111), where the coefficients (8.110) are

$$h_n = \frac{1}{2} A i^n \int_{-\infty}^{\infty} J_1(\Omega)\, e^{-\sigma^2\Omega^2/2}\, \Omega^{n-1}\, d\Omega \tag{8.115}$$

Since $J_1(\Omega)$ is odd, the integral in (8.110) vanishes unless n is even, i.e., $n = 2m$. Introducing the notation

$$g_{pq}(1/\sigma) = \sigma^q \int_0^{\infty} J_p(\Omega)\, e^{-\sigma^2\Omega^2/2}\, \Omega^{q-1}\, d\Omega = \int_0^{\infty} J_p(x/\sigma)\, e^{-x^2/2}\, x^{q-1}\, dx\,, \tag{8.116}$$

we have

$$h_n = \begin{cases} 0 & \text{for odd } n\,, \\ \dfrac{Ai^n}{\sigma^n} g_{1n}(1/\sigma) & \text{for } n = 2m\,. \end{cases} \tag{8.117}$$

Then, substituting (8.117) into (8.111), we obtain

$$k_\eta(\tau) = A^2 \sum_{m=1}^{\infty} \frac{1}{(2m)!}\, g_{1.2m}^2(1/\sigma)\, R^{2m}(\tau)\,.$$

The coefficients g_{pq} can be expressed in terms of the confluent hypergeometric function by using formula (4.434.2) of the handbook just quoted,[7] i.e.,

$$\begin{aligned} g_{pq}(\alpha) &= \int_0^{\infty} J_p(\alpha x)\, e^{-x^2/2}\, x^{q-1}\, dx \\ &= 2^{-1+(p-q)/2} \frac{\Gamma[(p+q)/2]}{\Gamma(p+1)}\, \alpha^p e^{-\alpha^2/2}\, {}_1F_1\left(\frac{p+q}{2} + 1; p+1; \frac{\alpha^2}{2}\right), \end{aligned} \tag{8.118}$$

where $p + q > 0$.

[7] I. M. Ryshik and I. S. Gradstein, *op. cit.*, p. 237.

4.2. Next, we consider the case where the input process is the sum of a Gaussian process $\xi(t)$ with the characteristic functions (8.102) and a "sine wave"

$$s(t) = E\cos(\omega_0 t + \varphi_0) = E\cos\Phi,$$

with fixed amplitude E and completely random initial phase φ_0 (cf. p. 198) which is statistically independent of $\xi(t)$. Then the characteristic function of the total signal

$$x(t) = \xi(t) + s(t) \tag{8.119}$$

is just the product of the characteristic functions of the processes $\xi(t)$ and $s(t)$:

$$\langle e^{i\Omega x}\rangle = \langle e^{i\Omega\xi}\rangle\,\langle e^{iE\Omega\cos\Phi}\rangle = e^{-\sigma^2\Omega^2/2}\,J_0(E\Omega). \tag{8.120}$$

Here, to calculate $\langle e^{iE\Omega\cos\Phi}\rangle$, we have averaged over the random initial phase φ_0, using the integral representation of the Bessel function $J_0(z)$. Similarly, for the two-dimensional characteristic function, we have

$$\begin{aligned}\langle e^{i\Omega x + i\Omega_\tau x_\tau}\rangle &= \langle\exp\{i\Omega\xi + i\Omega_\tau\xi_\tau\}\rangle\,\langle\exp\{iE(\Omega\cos\Phi + \Omega_\tau\cos\Phi_\tau)\}\rangle \\ &= \exp\left\{-\frac{1}{2}\sigma^2(\Omega^2 + 2R\Omega\Omega_\tau + \Omega_\tau^2)\right\} J_0\left(E\sqrt{\Omega^2 + 2\Omega\Omega_\tau\cos\omega_0\tau + \Omega_\tau^2}\right)\end{aligned} \tag{8.121}$$

[cf. (7.82)].

To determine the mean values $\langle\eta\rangle$ and $\langle\eta\eta_\tau\rangle$, we substitute the characteristic functions (8.120) and (8.121) [instead of the functions (8.102)] into (8.106). In calculating $\langle\eta\eta_\tau\rangle$, it is convenient to use the familiar addition theorem for Bessel functions[8]

$$J_0\left(E\sqrt{\Omega^2 + 2\Omega\Omega_\tau\cos\omega_0\tau + \Omega_\tau^2}\right) = \sum_{k=0}^{\infty}(-1)^k\,\epsilon_k J_k(E\Omega)\,J_k(E\Omega_\tau)\cos k\omega_0\tau, \tag{8.122}$$

[8] I. M. Ryshik and I. Ş. Gradstein, *op. cit.*, formula (6.541.1), p. 335.

where

$$\epsilon_0 = 1\,, \qquad \epsilon_1 = \epsilon_2 = \ldots = 2\,.$$

We also use the expansion (8.108) for the quantity $\exp(-\sigma^2 R\Omega\Omega_\tau)$ appearing in (8.121). After making all these substitutions, we obtain the following double series for the moment $\langle \eta\eta_\tau \rangle$ of the output signal:

$$\langle \eta\eta_\tau \rangle = \sum_{n=0}^{\infty}\sum_{k=0}^{\infty}\left[\frac{\sigma^n}{2\pi}\int_L F(i\Omega)\, J_k(E\Omega)\, e^{-\sigma^2\Omega^2/2}\,\Omega^n\, d\Omega\right]^2 \times \frac{(-1)^{n+k}}{n!}\,\epsilon_k R^n \cos k\omega_0\tau\,. \tag{8.123}$$

To obtain the correlation function $k_\eta(\tau)$, we have to subtract from (8.123) the term corresponding to $n = 0$, $k = 0$, which equals $\langle \eta \rangle^2$.

Example 2. Suppose the process (8.119) undergoes the transformation

$$\eta = g(x) = \begin{cases} x & \text{for} \quad x > 0\,, \\ 0 & \text{for} \quad x < 0\,. \end{cases} \tag{8.124}$$

Then we have

$$F(i\Omega) = \int_0^\infty e^{-i\Omega x}\, x\, dx = -\frac{1}{\Omega^2}\,, \tag{8.125}$$

where the contour of integration L coincides with the real axis $\operatorname{Im}\Omega = 0$ when the integral has no singularity at the origin $\Omega = 0$, and otherwise has a downward indentation avoiding the singularity at the origin. According to (8.123) and (8.125), we have

$$k_\eta(\tau) = \frac{\sigma^2}{4\pi^2}\sum_{\substack{n,k=0 \\ n+k>0}}\left[\sigma^{n-1}\int_L J_k(E\Omega)\, e^{-\sigma^2\Omega^2/2}\,\Omega^{n-2}\, d\Omega\right]^2 \times \frac{(-1)^{n+k}}{n!}\,\epsilon_k R^n(\tau)\cos k\omega_0\tau\,. \tag{8.126}$$

Only two integrands in this sum have a pole at the origin, i.e., those

corresponding to $n = 1$, $k = 0$ and $n = 0$, $k = 1$. Since these integrands are odd functions of Ω, it is easily seen that

$$\int_L J_0(E\Omega)\, e^{-\sigma^2\Omega^2/2}\, \frac{d\Omega}{\Omega} = i\pi\,,$$
$$\frac{1}{\sigma}\int_L J_1(E\Omega)\, e^{-\sigma^2\Omega^2/2}\, \frac{d\Omega}{\Omega^2} = \frac{E}{2\sigma}\, i\pi\,. \tag{8.127}$$

All the other integrals in (8.126) can be expressed in terms of th e functions (8.118). As a result, (8.126) reduces to

$$k_\eta(\tau) = \frac{\sigma^2}{4} R(\tau) + \frac{E^2}{8}\cos\omega_0\tau + \frac{\sigma^2}{4\pi^2}\sum_{\substack{n,k=0\\ n+k>1}} \frac{1}{n!}\,[1 + (-1)^{n+k}]^2\, \epsilon_k g^2_{k,n-1}(E/\sigma)\, R^n(\tau)\cos k\omega_0\tau\,. \tag{8.128}$$

In particular, (8.128) is valid for a narrow-band random process

$$\xi(t) = A(t)\cos(\omega_0 t + \varphi)\,,$$

with the correlation function

$$R(\tau) = r(\tau)\cos\omega_0\tau\,, \tag{8.129}$$

where $A(t)$, $\varphi(t)$ and $r(\tau)$ are slowly varying functions, e.g., $r'_\tau/r \ll \omega_0$. In this case, the output signal

$$\eta(t) = g[A\cos(\omega_0 t + \varphi) + E\cos(\omega_0 t + \varphi_0)] = g[B\cos(\omega_0 t + \psi)] \tag{8.130}$$

has spectral components lying in narrow bands near the frequencies $l\omega_0$ ($l = 0, 1, 2, \ldots$), and the correlation function (8.128) can be written as a series

$$k_\eta(\tau) = k_0(\tau) + k_1(\tau)\cos\omega_0\tau + k_2(\tau)\cos 2\omega_0\tau + \ldots, \tag{8.131}$$

involving the slowly varying functions $k_0(\tau)$, $k_1(\tau)$, $k_2(\tau)$, ...

Of particular interest is the first function $k_0(\tau)$, obtained by averaging $k_\eta(\tau)$:

$$k_0(\tau) = \frac{\omega_0}{2\pi} \int_\tau^{\tau+(2\pi/\omega_0)} k_\eta(\tau')\, d\tau' . \tag{8.132}$$

We now show that in the case of narrow-band signals, the averaged correlation function (8.132) is approximately equal to the correlation function $\mathbf{K}[\bar{\eta}, \bar{\eta}_\tau]$ of the time-averaged signal

$$\bar{\eta}(t) = \frac{\omega_0}{2\pi} \int_t^{t+(2\pi/\omega_0)} \eta(t')\, dt' = \frac{1}{T_0} \int_0^{T_0} \eta(t+x)\, dx , \tag{8.133}$$

where $T_0 = 2\pi/\omega_0$. According to (8.133), we have

$$\mathbf{K}[\bar{\eta}, \bar{\eta}_\tau] = \frac{1}{T_0^2} \int_0^{T_0} \int_0^{T_0} k_\eta(\tau + x - y)\, dx\, dy . \tag{8.134}$$

If we substitute (8.131) into (8.134), using the Taylor's series

$$k_\eta(\tau + x - y) = \sum_{r=0}^{\infty} \frac{1}{r!} k_s^{(r)}(\tau)\, (x-y)^r \qquad (r = 0, 1, 2, \ldots) ,$$

then, since

$$\int_0^{T_0} \int_0^{T_0} k_s(\tau) \cos k\omega_0(\tau + x - y)\, dx\, dy = 0 \quad \text{for} \quad k > 0$$

and

$$\left| \int_0^{T_0} \int_0^{T_0} (x-y)^r \cos k\omega_0(\tau + x - y)\, dx\, dy \right|$$

$$< \int_0^{T_0} \int_0^{T_0} |x-y|^r\, dx\, dy = \frac{2T_0^{r+2}}{(r+1)(r+2)} ,$$

we find that

$$\mathbf{K}[\bar{\eta}, \bar{\eta}_\tau] = k_0(\tau) + O(k_0' T_0 + k_1' T_0 + \ldots) .$$

It follows that

$$k_0(\tau) \approx \mathbf{K}[\bar{\eta}, \bar{\eta}_\tau] ,$$

as asserted, since the narrow-band condition implies $k_s' T_0 \ll k_s$.

It is most convenient to carry out the averaging (8.132) of the expression (8.123) before the expansion (8.108) in powers of R is made in (8.121). Thus, expanding only the Bessel function factor in (8.121), and using (8.129), we find that

$$\langle e^{i\Omega x + i\Omega_\tau x_\tau}\rangle = e^{-\sigma^2(\Omega^2+\Omega_\tau^2)/2}\sum_{k=0}^{\infty}(-1)^k\,\epsilon_k J_k(E\Omega)\,J_k(E\Omega_\tau)\,e^{-\sigma^2\Omega\Omega_\tau r\cos\omega_0\tau}\cos k\omega_0\tau. \tag{8.135}$$

We then carry out an average of the type (8.132) with respect to τ, regarding $r(\tau)$ as constant; as a result, (8.135) becomes

$$e^{-\sigma^2(\Omega^2+\Omega_\tau^2)/2}\sum_{k=0}^{\infty}(-1)^k\,\epsilon_k J_k(E\Omega)\,J_k(E\Omega_\tau)\,I_k(-\sigma^2 r\Omega\Omega_\tau), \tag{8.136}$$

where $I_k(z)$ is the modified Bessel function of the first kind, of order k.[9] Since

$$I_k(-\sigma^2 r\Omega\Omega_\tau) = \sum_{n=0}^{\infty}\frac{1}{n!(n+k)!}\left(-\frac{1}{2}\sigma^2 r\Omega\Omega_\tau\right)^{2n+k}$$

by analogy with (8.123) we have

$$\langle\bar\eta\bar\eta_\tau\rangle = \sum_{n,k=0}^{\infty}\left[\frac{\sigma^{2n+k}}{2\pi}\int_L F(i\Omega)\,J_k(E\Omega)\,e^{-\sigma^2\Omega^2/2}\,\Omega^{2n+k}\,d\Omega\right]^2 \times\frac{\epsilon_k}{n!(n+k)!}\left[\frac{1}{2}r(\tau)\right]^{2n+k}. \tag{8.137}$$

Subtracting $\langle\bar\eta\rangle^2 = \langle\eta\rangle^2$, the square of the mean value, from (8.137), we obtain the correlation function

$$k_{\bar\eta}(\tau) = k_0(\tau) = \sum_{\substack{n,k=0\\ n+k>0}}\left[\frac{\sigma^{2n+k}}{2\pi}\int_L F(i\Omega)\,J_k(E\Omega)\,e^{-\sigma^2\Omega^2/2}\,\Omega^{2n+k}\,d\Omega\right]^2\frac{\epsilon_k}{n!(n+k)!}\left[\frac{1}{2}r(\tau)\right]^{2n+k}, \tag{8.138}$$

since $\langle\eta\rangle^2$ is just the term in (8.137) with $k = 0, n = 0$.

[9] I. M. Ryshik and I. S. Gradstein, *op. cit.*, formula (6.443), p. 317.

If the nonlinear transformation $g(x)$ has the form (8.124), then the d-c component (8.133) is proportional to the total amplitude, i.e.,

$$\bar{\eta} = \frac{1}{2\pi}\int_0^{2\pi} g[B\cos(\chi+\psi)]\,d\chi = \frac{B}{\pi}, \tag{8.139}$$

and hence (8.138) can be used to find the correlation function of the amplitude $B(t)$. Substituting (8.125) into (8.138), we find that

$$\begin{aligned} k_{\bar{\eta}}(\tau) &= \pi^{-2}k_B(\tau) \\ &= \frac{\sigma^2}{4\pi^2}\sum_{\substack{n,k=0\\ n+k>0}}\left[\sigma^{2n+k-1}\int_{-\infty}^{\infty} J_k(E\Omega)\,e^{-\sigma^2\Omega^2/2}\,\Omega^{2n+k-2}\,d\Omega\right]^2 \\ &\quad\times\frac{\epsilon_k}{n!(n+k)!}\left[\frac{1}{2}r(\tau)\right]^{2n+k}, \end{aligned}$$

so that, according to (8.118),

$$k_B(\tau) = \sigma^2\sum_{\substack{n,k=0\\ n+k>0}}\frac{\epsilon_k}{n!(n+k)!2^{2n+k}}g^2_{k,2n+k-1}(E/\sigma)\,r^{2n+k}(\tau). \tag{8.140}$$

CHAPTER 9

Nonlinear Transformations with Memory. Detection of Random Signals

To illustrate the basic idea of the methods used to study what happens when random signals undergo nonlinear transformations with memory, we shall consider the example of the *diode detector*, whose action is described by a first-order nonlinear differential equation. By the diode detector, we mean the device shown in Figure 5, obtained by connecting a diode, i.e., a nonlinear element with current-voltage characteristic $I = F(V)$, to a parallel RC-circuit.[1] Suppose we want to find the statistical characteristics of the voltage $\eta(t)$ across the RC-circuit, when the input to the detector is a random signal $\xi(t)$ with known properties. The differential equation for the diode detector has the form

$$\dot{\eta} + \frac{1}{RC}\eta = \frac{1}{C}F(\xi - \eta) \tag{9.1}$$

[cf. (8.3)]. A rigorous solution of this nonlinear differential equation for typical functions F approximating actual current-voltage characteristics is impossible even in the case where the function $\xi(t)$ in the right-hand side of (9.1) is an ordinary (non-constant) function of time, rather than a random function. Thus, solving (9.1) is certainly all the more unfeasible when $\xi(t)$ is random, and we

[1] The ordinary diode rectifier, with a purely resistive load, is a simpler system, since there is no need to take the reaction of the load into account. In this case, the problem reduces to the study of a zero-memory transformation of the input process, like those studied in the previous chapter.

must have recourse to approximate methods of solution. The choice of one method as opposed to another depends on the form of the function $F(V)$, on the system parameters, and also on the relation between the correlation time τ_{cor} of the input process and the time constant of the detector (equal to RC). We can distinguish five special cases, each of which corresponds to a different way of handling the problem:

1. If the inequality

$$\tau_{cor} \gg RC \tag{9.2}$$

holds, we can study the problem in the *quasi-static approximation.* In this case, as a first approximation we can neglect the time derivative $\dot{\eta}$ in equation (9.1), and then the problem reduces to the zero-memory nonlinear transformation

$$\eta = RF(\xi - \eta).$$

2. Suppose the input process $\xi(t)$ is narrow-band, i.e., corresponds to oscillations of frequency ω_0 with slowly varying amplitude and phase, so that $\xi(t)$ has a large correlation time τ_{cor}. If the relation

$$\tau_{cor} \gg RC \gg \frac{2\pi}{\omega_0} \tag{9.3}$$

holds, then we can apply the envelope method, which is a variant of the quasi-static method. This case includes the detector stages used in many radio receivers.

3. In many cases, it is appropriate to use the method of "small nonlinearity," i.e., the problem is regarded as linear in the first approximation, and the nonlinear effects are only taken into account in subsequent approximations. This method is applicable when the nonlinear term (or terms) in the original equation is small, or when the equation can be linearized with respect to deviations from some suitably chosen zeroth approximation. The application of this method is not limited to the case where there is some special relation between the time constant of the system and the correlation time, and in fact the method can even be applied when $\tau_{cor} \sim RC$.

4. If the time constant is small, which means that

$$\tau_{cor} \ll RC, \tag{9.4}$$

then $\xi(t)$ can be regarded as a Markov process, and we can make use of the Fokker-Planck equation. In particular, this case may arise in making measurements, when we want to ascertain the mean value $\langle\eta\rangle$ of the output signal (characterizing the reading of the measuring device) and the variance $\mathbf{D}\eta$ (characterizing the measurement error). Then, it is sufficient to know the one-dimensional stationary distribution, which is a solution of the Fokker-Planck equation.

5. When we are dealing with intermediate values of the correlation time, i.e.,

$$\tau_{cor} \sim RC,$$

and when the signals are strong, the problem becomes very difficult. If the input signal $\xi(t)$ is assumed to be a Markov process, the output voltage will be a component of a multidimensional Markov process, and some results can be obtained by using the multidimensional Fokker-Planck equation.

We now study these various methods in turn.

1. Detection of Narrow-Band Processes. The Envelope Method

If the input process is slowly varying ($\tau_{cor} \gg RC$), then, as already noted, we can neglect the derivative $\dot{\eta}$ in equation (9.1), obtaining

$$\eta = RF(\xi - \eta). \tag{9.5}$$

Solving (9.5) for η in terms of ξ (for example, graphically), we obtain a zero-memory transformation

$$\eta(t) = g[\xi(t)]. \tag{9.6}$$

Methods for analyzing nonlinear transformations of this kind were

presented in Chap. 8. It is sometimes convenient to represent the function (9.6) as a polynomial, e.g., of the form

$$g(\xi) = a_1\xi + a_2\xi^2,$$

or

$$g(\xi) = a_2\xi^2 - a_4\xi^4.$$

Then, the coefficients a_1, a_2 and a_4 are chosen numerically in such a way that the polynomials give a satisfactory approximation over the important part of the transfer characteristic. The statistical characteristics of the output signal for these two polynomial transfer characteristics were calculated in Examples 1 and 2 of Chap. 8, Sec. 2.

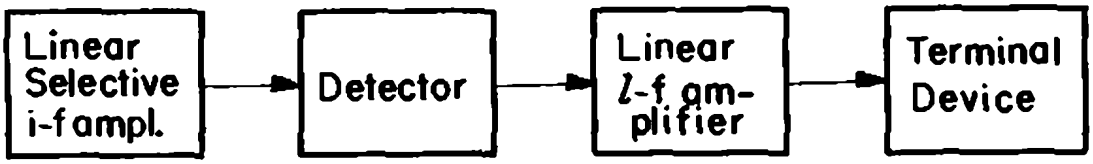

Fig. 11. Simplified block diagram of a typical radio receiver.

We now examine in detail a variant of the quasi-static method which is suitable in the case where the input signal is a narrow-band process. Consider a typical radio receiver (shown schematically in Figure 11), consisting of a linear selective intermediate-frequency (i-f) amplifier, a detector stage, a linear low-frequency (l-f) amplifier, and a terminal device. Usually, the i-f amplifier is a frequency-selective four-terminal device, i.e., the inequality $\Delta\omega \ll \omega_0$ usually holds, where ω_0 is some "center" (angular) frequency, and $\Delta\omega$ is the width of the pass band of the amplifier. Thus, if the input to the i-f amplifier is a broad-band stationary process, the output $\eta(t)$ is a narrow-band process, which usually has a correlation function of the form

$$R(\tau) = r(\tau)\cos\omega_0\tau,$$

where $r(\tau)$ is slowly varying compared to $\cos\omega_0\tau$. [In formula (7.4), the function $s(\tau)$ vanishes because of the symmetry of the frequency response of the i-f amplifier about the center frequency ω_0.]

Like every narrow-band process, $\xi(t)$ can be written as an oscillation

$$\xi(t) = A(t)\cos[\omega_0 t + \varphi(t)], \tag{9.7}$$

with slowly varying amplitude and phase (7.55). If in addition to the process $\xi(t)$, the input to the detector contains a sine wave $E\cos\omega_0 t$, then the total signal $\xi(t) + E\cos\omega_0 t$ can also be written as an oscillation of the form (7.83), with a different amplitude $B(t)$ and phase $\psi(t)$. Thus, the results to be obtained below can be extended to this case, if we simply replace the amplitude $A(t)$ by $B(t)$.

The function of any detector is to "extract" the low-frequency modulation from its input in as effective a manner as possible, and hence the output voltage across the RC-circuit must manage to "follow" the amplitude changes. Since the amplitude varies with the time constant τ_{cor}, this means that the condition

$$\tau_{cor} \gg RC \tag{9.8}$$

must hold. On the other hand, the detector must exert a strong smoothing action on the radio-frequency (r-f) oscillations, which means that the condition

$$RC \gg \frac{2\pi}{\omega_0} \tag{9.9}$$

must hold. If these two conditions are satisfied, the study of the detection problem is greatly simplified, since then, as we shall see, the output voltage $\eta(t)$ is related to the amplitude $A(t)$ [or $B(t)$] of the input oscillation by a zero-memory transformation.

Substituting (9.7) into (9.1), we obtain

$$\dot{\eta} + \frac{1}{RC}\eta = \frac{1}{C}F[A\cos(\omega_0 t + \varphi) - \eta]. \tag{9.10}$$

Integrating (9.10) over a period from t to $t + (2\pi/\omega_0)$, we find that

$$\eta\left(t + \frac{2\pi}{\omega_0}\right) - \eta(t) \tag{9.11}$$
$$= \frac{1}{RC}\int_t^{t+(2\pi/\omega_0)} \{-\eta(t') + RF[A(t')\cos(\omega_0 t' + \varphi) - \eta(t')]\}\, dt'.$$

However, because of (9.9), the function $\eta(t)$ only manages to change very slightly during one period. Therefore, the difference

$$\eta\left(t + \frac{2\pi}{\omega_0}\right) - \eta(t)$$

is almost equal to $(2\pi/\omega_0)\dot{\eta}$, and $\eta(t')$ can be regarded as constant in the integrand, i.e., $\eta(t') \approx \eta(t)$. The amplitude A and the phase φ can also be regarded as constant during a period, so that we need only take into account the change of $\cos[\omega_0 t' + \varphi(t)]$. Thus, (9.11) reduces to

$$\dot{\eta} + \frac{1}{RC}\eta = \frac{1}{C}\frac{\omega_0}{2\pi}\int_t^{t+(2\pi/\omega_0)} F[A\cos(\omega_0 t' + \varphi) - \eta]\, dt$$

or

$$\dot{\eta} + \frac{1}{RC}\eta = \frac{1}{C}\frac{1}{2\pi}\int_0^{2\pi} F[A(t)\cos\chi - \eta]\, d\chi\,. \tag{9.12}$$

The right-hand side is essentially a known function of A and η, obtained by integrating the current-voltage characteristic $F(V)$ of the diode.

As it stands, equation (9.12) is hard to solve in general form. However, the problem is greatly simplified by the condition (9.8) which requires that the amplitude vary slowly; in this case, we can neglect the derivative $\dot{\eta}$ in (9.12). In fact, just as in the case of $A(t)$ itself, the only important spectral components of the process $\eta(t)$ lie in the frequency band

$$|\omega - \omega_0| \sim \Delta\omega = \frac{1}{\tau_{cor}}.$$

For these frequencies, the derivative $\dot{\eta}_\omega = i\omega\eta_\omega$ is much smaller than η_ω/RC, because of (9.8). Therefore, the derivative $\dot{\eta}$ in (9.12) is much smaller than the second term η/RC. If we neglect $\dot{\eta}$, (9.12) becomes

$$\eta = \frac{R}{2\pi}\int_0^{2\pi} F(A\cos\chi - \eta)\, d\chi\,, \tag{9.13}$$

which specifies the zero-memory transformation leading from the amplitude $A(t)$ to the output voltage $\eta(t)$.

We now use (9.13) to study three typical cases of diode detection, where the current-voltage characteristic $F(V)$ is piecewise linear, quadratic and exponential, respectively.

Example 1. In this case, usually called the *linear detector*, the current-voltage characteristic is

$$I = F(V) = \begin{cases} \frac{1}{R_i} V & \text{for} \quad V > 0, \\ 0 & \text{for} \quad V < 0, \end{cases} \tag{9.14}$$

where R_i is the internal resistance of the diode in its conducting state. Substituting (9.14) into (9.13), we have

$$\eta = \frac{AR}{\pi R_i} \int_0^\gamma \left(\cos \chi - \frac{\eta}{A}\right) d\chi, \tag{9.15}$$

where

$$\gamma = \arccos \frac{\eta}{A}.$$

We now introduce the dimensionless quantity

$$k = \frac{\eta}{A}, \tag{9.16}$$

which can be called the "envelope reproduction coefficient." Then (9.15) becomes

$$k = \frac{R}{\pi R_i} \int_0^{\arccos k} (\cos \chi - k)\, d\chi,$$

and carrying out the integration, we finally obtain

$$\frac{R}{R_i} = \frac{\pi k}{\sqrt{1 - k^2} - k \arccos k}. \tag{9.17}$$

If $R/R_i \lesssim 1$, the envelope reproduction coefficient k is < 0.3, and then (9.17) implies that

$$k \approx \frac{1}{\pi} \frac{R}{R_i}.$$

However, such values of R/R_i are not encountered in typical radio receivers. In fact, this ratio is of the order of 3 to 10 in radar and television receivers, and can equal several hundred in communication receivers. In Figure 12, we plot the coefficient k for large values of R/R_i. It is clear from the figure that the envelope reproduction coefficient k grows monotonically with R/R_i, and approaches 1 asymptotically.

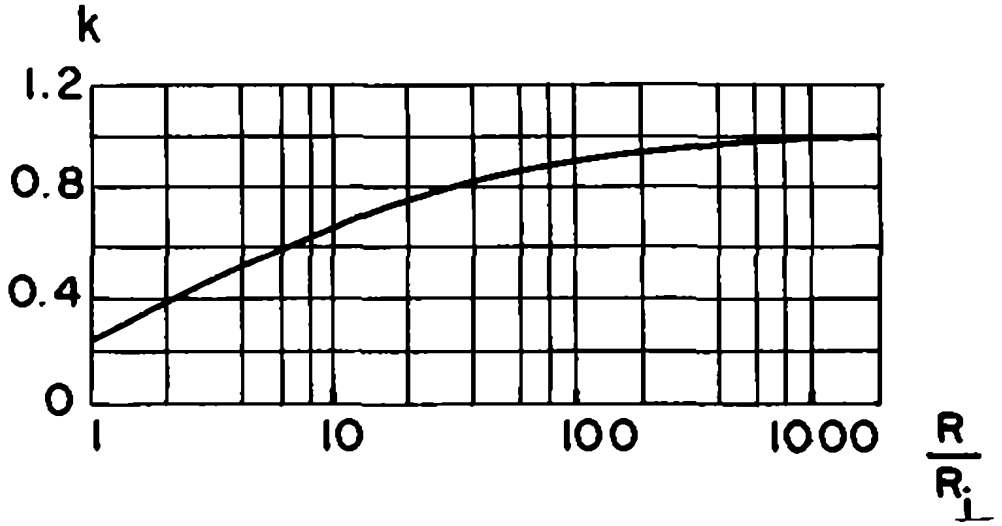

Fig. 12. Dependence of the envelope reproduction coefficient k on the resistance ratio R/R_i for the linear detector.

The characteristic feature of the linear detector, as opposed to other kinds of detectors, is that its output voltage $\eta(t)$ depends linearly on $A(t)$, and therefore it is appropriate to introduce the envelope reproduction coefficient (9.16), which does not depend on A. After calculating k, the statistical characteristics of the output signal

$$\eta(t) = kA(t) \tag{9.18}$$

can be found very simply from the corresponding characteristics of $A(t)$. In particular, we have

$$\langle \eta \rangle = k\langle A \rangle, \qquad \langle \eta\eta_\tau \rangle = k^2\langle AA_\tau \rangle,$$

where it is useful to recall that $\langle AA_\tau \rangle$ is given by the third of the formulas (8.56).

Example 2. For the *quadratic detector*, the current-voltage characteristic is given by

$$F(V) = \begin{cases} \beta V^2 & \text{for} \quad V > 0, \\ 0 & \text{for} \quad V < 0, \end{cases} \tag{9.19}$$

where the coefficient β is a dimensionless constant. Using (9.13), we obtain

$$\eta = \beta \frac{RA^2}{\pi} \int_0^{\gamma} \left(\cos \chi - \frac{\eta}{A}\right)^2 d\chi ,$$

where

$$\gamma = \arccos \frac{\eta}{A} ,$$

as before; again introducing $k = \eta/A$, we have

$$k = \beta \frac{RA}{\pi} \int_0^{\arccos k} (\cos \chi - k)^2 \, d\chi .$$

Carrying out the integration, we obtain the following transcendental equation, after making some elementary calculations:

$$\beta RA = \frac{2\pi k}{(1 + 2k^2) \arccos k - 3k \sqrt{1 - k^2}} . \tag{9.20}$$

In this case, the envelope reproduction coefficient k no longer has its former meaning, but depends on A.

If $\beta RA \lesssim 0.1$, the inequality $k \ll 1$ holds, and then, setting $\arccos k \approx \pi/2$, we find from (9.20) that

$$k \approx \tfrac{1}{4} \beta RA , \qquad \eta \approx \tfrac{1}{4} \beta RA^2 . \tag{9.21}$$

For large values of βRA, the coefficient k can be found by solving equation (9.20) numerically. In Figure 13, we plot the function

$$k = f(\beta RA) , \tag{9.22}$$

where $f(x)$ is the inverse of the function

$$x = \frac{2\pi k}{(1 + 2k^2) \arccos k - 3k \sqrt{1 - k^2}} .$$

As the figure shows, the coefficient k grows monotonically with βRA, and approaches 1 asymptotically.

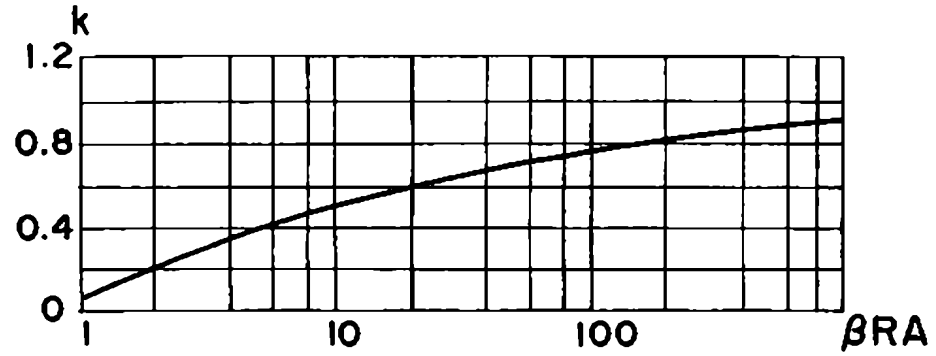

Fig. 13. Dependence of the coefficient k on the parameter βRA for the quadratic detector.

It follows from (9.22) that the dependence of the output voltage on the amplitude can be written in the form

$$\beta R_\eta = \beta RA\, f(\beta RA)\,. \tag{9.23}$$

The function (9.23) can be approximated by a polynomial in some region or other, depending on the range of values of βRA most frequently encountered. For example, the quadratic polynomial

$$\beta R_\eta = 0.583(\beta RA) + 0.002(\beta RA)^2$$

plotted in Figure 14 gives a satisfactory approximation to (9.23)

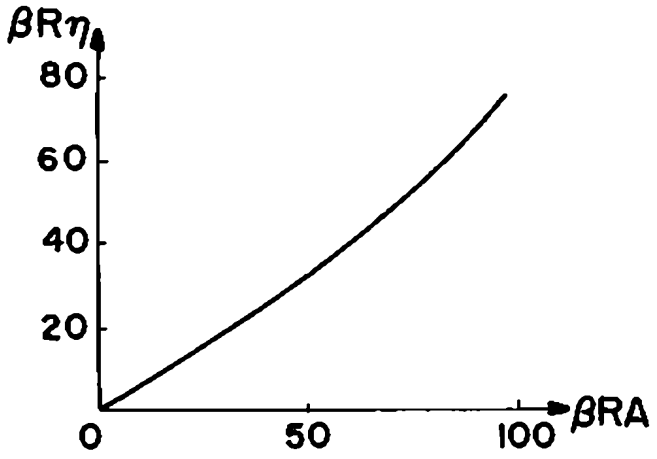

Fig. 14. Approximate dependence of the output voltage on the amplitude for the quadratic detector.

in the range $\beta RA = 0$ to 100, but it is best to use this approximation when $\beta RA \sim 50$, since it is not too accurate for $\beta RA \ll 50$. For amplitudes corresponding to a different range of βRA, one should use a polynomial with different coefficients. After choosing the appropriate polynomial, the statistical characteristics of the output signal can be calculated by using the method given in Example 2 of Chap. 8, Sec. 2.

Example 3. Finally, we consider the *exponential detector*. In this case, the current-voltage characteristic has the exponential form

$$I = F(V) = i_0 e^{aV}, \tag{9.24}$$

where i_0 and a are constants, I is the current through the nonlinear element, and V is the voltage across it. The curve (9.24) can be used to approximate the current-voltage characteristics of actual diodes. with

$$i_0 = 50 \text{ to } 600 \text{ microamperes},$$

$$a = 5 \text{ to } 10 \text{ volts}^{-1},$$

in the case of oxide-coated cathodes at normal working temperatures.

If the input to the detector is the process (9.7), then, according to (9.13), the voltage across the RC-circuit is given by

$$\eta = Ri_0 \frac{1}{2\pi} \int_0^{2\pi} \exp\{a(A\cos\chi - \eta)\}\, d\chi = Ri_0 e^{-a\eta} I_0(aA)$$

or

$$a\eta e^{a\eta} = aRi_0 I_0(aA). \tag{9.25}$$

Here $I_0(z) = J_0(iz)$ is the modified Bessel function of the first kind, of order zero. Equation (9.25) can be solved in various ways, depending on whether the quantity $aRI_0(a\langle A\rangle)$ or the quantity $a\langle A\rangle$ is large or small compared to unity.

If $a\langle A\rangle \ll 1$, then η differs very slightly from the constant voltage η_0 which appears in the absence of any input voltage, and corresponds to setting $A = 0$ in (9.25):

$$a\eta_0 e^{a\eta_0} = aRi_0. \tag{9.26}$$

Using (9.26) and the expansions

$$I_0(aA) = 1 + \frac{1}{4}(aA)^2 + \frac{1}{64}(aA)^4 + \dots ,$$

$$xe^x = x_0e^{x_0} + (1 + x_0)\, e^{x_0}(x - x_0) + (x - x_0)^2 ,$$

we find from (9.25) that

$$\eta - \eta_0 \approx \frac{a\eta_0}{4(1 + a\eta_0)}\, aA^2 . \tag{9.27}$$

Therefore, for small values $aA \lesssim 0.1$, the voltage across the detector load due to the presence of the input signal is proportional to the square of the envelope. In fact, we have quadratic detection quite generally for small input amplitudes, regardless of the form of the current-voltage characteristic $F(V)$. This is related to the fact that the output signal is an even function of A, since it does not depend on the phase of the input signal, and does not change when the phase is shifted by 180°, i.e., when A is replaced by $-A$. Thus, only even powers of A appear in the power series expansion of $\eta - \eta_0$, and hence for small values of A, only the first term of the series, proportional to A^2, is important.

The case where ARi_0 and aA are both large, i.e.,

$$aRi_0I_0(aA) \gg 1 , \tag{9.28}$$

is of great practical interest. Then the envelope reproduction coefficient η/A is large, of order unity. Using the condition (9.28), we can solve equation (9.25) by the method of successive approximations. Writing (9.25) in the form

$$x + \ln x = z ,$$

where

$$x = a\eta , \qquad z = \ln [aRi_0I_0(aA)] ,$$

we define the successive approximations by

$$x_n = z - \ln x_{n-1} \qquad (n = 1, 2, \dots) ,$$

i.e.,

$$x_1 = z,$$
$$x_2 = z - \ln z,$$
$$x_3 = z - \ln(z - \ln z) \approx z - \ln z + \frac{\ln z}{z},$$
$$\cdots\cdots\cdots$$

Confining ourselves to the second approximation, we have

$$\eta = \frac{1}{a} \ln \frac{aRi_0 I_0(aA)}{\ln [aRi_0 I_0(aA)]}. \tag{9.29}$$

Next, we use (9.29) to determine the coefficients a_1 and a_2 in the expression (8.42), choosing c to be $\langle A \rangle$, the mean value of the amplitude:

$$\eta = a_0 + a_1(A - \langle A \rangle) + a_2(A - \langle A \rangle)^2.$$

Equating the terms in this polynomial to the corresponding terms of the Taylor's series expansion of (9.29), we obtain

$$a_0 = \frac{1}{a}(z_1 - \ln z_1),$$
$$a_1 = \left(1 - \frac{1}{z_1}\right)\vartheta, \tag{9.30}$$
$$a_2 = \frac{a}{2}\frac{\vartheta^2}{z_1^2} + \frac{a}{2}\left(1 - \frac{1}{z_1}\right)\left(1 - \frac{\vartheta}{a\langle A \rangle} - \vartheta^2\right),$$

where

$$z_1 = \ln aRi_0 + \ln I_0(a\langle A \rangle), \qquad \vartheta = \frac{I_1(a\langle A \rangle)}{I_0(a\langle A \rangle)}.$$

As we increase the parameter $a\langle A \rangle$,

$$z_1 \to \infty, \qquad \vartheta \to 1,$$

so that, according to (9.30), the coefficient $a_1 \to 1$, while $a_2 \to 0$. Thus, in this case, the transformation of the signal by the exponential detector approaches the ideal detector

$$\eta = a_0' + A \qquad (a_0' = a_0 - \langle A \rangle), \tag{9.31}$$

which reproduces the envelope exactly.

For an arbitrary current-voltage characteristic $F(V)$, the more steeply the function $F(V)$ grows compared to V/R, i.e., the more valid the condition

$$R_i \ll R, \quad \text{where} \quad R_i = \frac{1}{\partial F_i' \partial V} \quad \text{and} \quad V > 0, \tag{9.32}$$

the closer the detection is to being ideal.

2. The Method of Small Nonlinearity

We shall present the other methods of analyzing nonlinear transformations with memory by applying them to the special case where the current-voltage characteristic of the nonlinear element (diode) is an exponential function

$$F(V) = i_0 e^{aV}; \tag{9.33}$$

here, i_0 and a are the constants discussed earlier (p. 253). Then equation (9.1) becomes

$$\dot{\eta} + \frac{1}{RC}\eta = \frac{i_0}{C} e^{a\xi - a\eta}. \tag{9.34}$$

It is convenient to introduce the new variable

$$x = e^{a\eta}, \tag{9.35}$$

which satisfies the equation

$$RC\dot{x} + x \ln x = ai_0 R e^{a\xi}. \tag{9.36}$$

The character of the process $x(t)$ satisfying (9.36) is determined by the size of the parameter ai_0R, and also by the ratio of the correlation time τ_{cor} of the process $\xi(t)$ to the time constant RC. Depending on whether these parameters ai_0R and τ_{cor}/RC are large or small, we can apply various special asymptotic methods for studying $x(t)$. For example, the condition $\tau_{cor} \gg RC$ allows us to use the quasi-static method. On the other hand, if $\tau_{cor} \ll RC$, we can use stochastic methods involving the Fokker-Planck equation.

However, neither of these methods is suitable if $\tau_{cor} \sim RC$. This compels us to have recourse to another method, which is appropriate when the other parameter ai_0R satisfies certain conditions. For the time being, we postpone the precise formulation of these conditions, and instead explain the method itself, which we call the *method of small nonlinearity*. In this method, we look for successive approximations to the solution $x(t)$ of equation (9.36). The zeroth approximation $x_0(t)$ is defined as the solution of the equation

$$RC\dot{x}_0 + x_0 \ln x_0 = ai_0R\langle e^{a\xi}\rangle , \tag{9.37}$$

which is suggested by averaging (9.36). Assuming that $\xi(t)$ is a stationary process, we are interested in a stationary process $x(t)$ which began in the remote past. This means that a stationary value of x_0 has been established, which, according to (9.37), satisfies the equation

$$x_0 \ln x_0 = ai_0R\langle e^{a\xi}\rangle . \tag{9.38}$$

The average $\langle e^{a\xi}\rangle$ is the value of the characteristic function $\langle e^{iu\xi}\rangle$ for $u = -ia$. If $\xi(t)$ is a Gaussian process with zero mean, then

$$\langle e^{iu\xi}\rangle = e^{-\sigma^2u^2/2}, \qquad \langle e^{a\xi}\rangle = e^{a^2\sigma^2/2}, \tag{9.39}$$

and (9.38) becomes

$$x_0 \ln x_0 = ai_0Re^{a^2\sigma^2/2} . \tag{9.40}$$

Using (9.35), we find that

$$\eta_0 = \frac{1}{a} \ln x_0$$

sarisfies the equation

$$\eta_0 e^{a\eta_0} = i_0Re^{a^2\sigma^2/2} . \tag{9.41}$$

Next, we derive a differential equation for the deviation

$$z = x - x_0 . \tag{9.42}$$

Subtracting (9.37) from (9.36), we find that

$$RC\dot{z} + (x_0 + z) \ln (x_0 + z) - x_0 \ln x_0 = ai_0R(e^{a\xi} - \langle e^{a\xi}\rangle) . \tag{9.43}$$

If the function $x \ln x$ can be expanded in a Taylor's series

$$x \ln x = x_0 \ln x_0 + (\ln x_0 + 1) z + \frac{1}{2x_0} z^2 - \frac{1}{6x_0^2} z^3 + \dots , \tag{9.44}$$

then equation (9.43) becomes

$$RC\dot{z} + (\ln x_0 + 1) z = ai_0 R\zeta - \frac{z^2}{2x_0} + \frac{z^3}{6x_0^2} - \dots , \tag{9.45}$$

where

$$\zeta = e^{a\xi} - \langle e^{a\xi} \rangle . \tag{9.46}$$

This equation contains the nonlinear terms

$$\frac{z^2}{2x_0} - \frac{z^3}{6x_0^2} + \dots = f(x_0 + z) - f(x_0) - f'(x_0) z , \tag{9.47}$$

where

$$f(x) = x \ln x .$$

The use of the method of small nonlinearity is based on the assumption that these nonlinear terms have comparatively little effect, and can be neglected in a first approximation. Therefore, (9.45) leads to the following equation for the first approximation:

$$RC\dot{z}_1 + (\ln x_0 + 1) z_1 = ai_0 R\zeta . \tag{9.48}$$

To clarify how higher-order approximations are constructed, we introduce an auxiliary small parameter ϵ. Then, we write the solution (9.42) as an expansion

$$z = z_1 + \epsilon z_2 + \epsilon^2 z_3 + \dots , \tag{9.49}$$

and we equip the nonlinear terms with appropriate small parameters. Thus, writing (9.45) in the form

$$RC\dot{z} + (\ln x_0 + 1) z = ai_0 R\zeta - \frac{\epsilon}{2x_0} z^2 + \frac{\epsilon^2}{6x_0^2} z^3 - \dots ,$$

substituting from (9.49), and equating terms of the same order in ϵ, we find that the equation for the next approximation is

$$RC\dot{z}_2 + (\ln x_0 + 1) z_2 = -\frac{z_1^2}{2x_0} . \tag{9.50}$$

When it is inconvenient to make the Taylor's series expansion (9.47), or when the expansion does not exist, the whole nonlinear term (9.47) appearing in (9.45) can be multiplied by a small parameter. With this version of our method, the whole nonlinear function

$$f(x_0 + z_1) - f(x_0) - f'(x_0)\, z_1$$

will appear in the right-hand side of the equation for z_2. Once the method of constructing the various approximations is clear, we can get rid of the small parameter, by setting it equal to 1.

The right-hand sides of (9.48) and (9.50) involve only known functions or functions already found in lower-order approximations. Moreover, these equations are linear in the unknown functions, and can be solved without any particuler difficulty. Thus, the stationary solution of equation (9.48) is

$$z_1(t) = \frac{ai_0}{C} \int_{-\infty}^{t} \exp\left\{-\frac{\ln x_0 + 1}{RC}(t - t')\right\} \zeta(t')\, dt' . \qquad (9.51)$$

According to the definition (9.46), the process $\zeta(t)$ has mean value zero, and hence

$$\langle z_1 \rangle = 0 . \qquad (9.52)$$

We now find the correlation function and the variance of the first approximation $z_1(t)$. First, we calculate the correlation function of the process $\zeta(t)$, assuming that $\xi(t)$ is a stationary Gaussian process with mean value zero and correlation function $k(\tau) = \sigma^2 R(\tau)$. Because of (9.46), we have

$$\langle \zeta \zeta_\tau \rangle = \langle e^{a\xi + a\xi_\tau} \rangle - \langle e^{a\xi} \rangle^2 . \qquad (9.53)$$

The quantity $\langle e^{a\xi + a\xi_\tau} \rangle$ can be expressed in terms of the two-dimensional characteristic function, given by the familiar formula (8.102). Thus, evaluating $\Theta_2(-ia, -ia)$ and using (9.39), we obtain

$$\langle \zeta \zeta_\tau \rangle = e^{a^2\sigma^2}[e^{a^2\sigma^2 R(\tau)} - 1] . \qquad (9.54)$$

According to (9.51), the correlation function $z_1 z_{1\tau}$ equals

$$\langle z_1 z_{1\tau}\rangle = \left(\frac{ai_0}{C}\right)^2 \int_{-\infty}^{t}\int_{-\infty}^{t+\tau} e^{-\beta(t-t_1)-\beta(t+\tau-t_2)}\,\langle\zeta(t_1)\,\zeta(t_2)\rangle\,dt_1\,dt_2\,, \quad (9.55)$$

where

$$\beta = \frac{\ln x_0 + 1}{RC}\,.$$

Making the change of variables

$$\rho = t_2 - t_1 - \tau\,, \qquad s = \frac{t_1 + t_2}{2}\,,$$

we can write (9.55) as

$$\langle z_1 z_{1\tau}\rangle = \left(\frac{ai_0}{C}\right)^2 \int_{-\infty}^{\infty} d\rho \int_{-\infty}^{t+\frac{1}{2}\tau-\frac{1}{2}|\rho|} e^{-\beta(2t-2s+\tau)}\,\langle\zeta\zeta_{\rho+\tau}\rangle\,ds\,.$$

After evaluating the integral with respect to s, this becomes

$$\langle z_1 z_{1\tau}\rangle = \left(\frac{ai_0}{C}\right)^2 \frac{1}{2\beta}\int_{-\infty}^{\infty} e^{-\beta|\rho|}\,\langle\zeta\zeta_{\rho+\tau}\rangle\,d\rho\,. \quad (9.56)$$

In particular, setting $\tau = 0$, we find the variance

$$\langle z_1^2\rangle = \left(\frac{ai_0}{C}\right)^2 \frac{1}{\beta}\int_0^{\infty} e^{-\beta\rho}\langle\zeta\zeta_\rho\rangle\,d\rho\,. \quad (9.57)$$

Using (9.54), we obtain

$$\langle z_1^2\rangle = \frac{(ai_0R)^2}{\ln x_0 + 1}\,e^{a^2\sigma^2}\,\frac{1}{RC}\int_0^{\infty}\left(e^{a^2\sigma^2R(\tau)} - 1\right)e^{-\beta\tau}\,d\tau\,. \quad (9.58)$$

To obtain more exact results, we can substitute (9.51) into the right-hand side of (9.50), and then calculate a higher-order approximation.

The statistical characteristics of the output voltage

$$\eta = \frac{1}{a}\ln(x_0 + z)$$

can now be found by expanding η in a Taylor's series

$$\eta = \eta_0 + \frac{z}{ax_0} - \frac{z^2}{2ax_0^2} + \dots , \tag{9.59}$$

and then substituting the series (9.49) into (9.59). If we confine ourselves to results involving the first approximations (9.51), (9.56) and (9.57), we have

$$\begin{aligned} \langle \eta \rangle &= \eta_0 - \frac{1}{2ax_0^2} \langle z_1^2 \rangle , \\ \langle (\eta - \eta_0)(\eta_\tau - \eta_0) \rangle &= \frac{1}{a^2 x_0^2} \langle z_1 z_{1\tau} \rangle , \end{aligned} \tag{9.60}$$

and in particular,

$$\mathbf{D}\eta = \frac{1}{a^2 x_0^2} \langle z_1^2 \rangle . \tag{9.61}$$

To determine the right-hand sides of these expressions, we use formulas (9.56) to (9.58).

We now analyze the conditions under which the above method is applicable, for the most difficult case, i.e., where the process $\xi(t)$ has a correlation time of the order of the filter time constant ($\tau_{cor} \sim RC$). The method of small nonlinearity is applicable if the nonlinear terms in equation (9.45) have a small effect compared to the linear term $(\ln x_0 + 1)z$. Choosing the first of the nonlinear terms, we obtain the condition

$$|(\ln x_0 + 1) z| \gg \left| \frac{z^2}{2x_0} \right|$$

or

$$|z| \ll 2x_0(\ln x_0 + 1) \tag{9.62}$$

where for z we can take the first approximation (9.51). According to (9.38) and (9.40), the quantity $x_0(\ln x_0 + 1)$ differs only slightly from

$$x_0 \ln x_0 = ai_0 Re^{a^2\sigma^2/2} \sim ai_0 Re^{-a^2\sigma^2/2}\, \sigma(\zeta) , \tag{9.63}$$

where $\sigma(\zeta)$ is the standard deviation of ζ, since according to (9.54),

$$\sigma(\zeta) \sim e^{a^2\sigma^2} .$$

On the other hand, using (9.51) we can obtain the estimate

$$z_1 \lesssim \frac{ai_0}{C} \frac{RC}{\ln x_0 + 1} \zeta , \tag{9.64}$$

where the inequality sign corresponds to the case where the function $\zeta(t)$ changes rapidly, with time constant

$$\tau_{cor} \lesssim \frac{RC}{\ln x_0 + 1} .$$

In the opposite case, where $\tau_{cor} \sim RC$, the quantity z_1 is chiefly determined by the factor $e^{-\beta(t-t')}$ appearing in the integrand of (9.51), and then the sign $\sim$ should be chosen in (9.64).

Substituting (9.64) into (9.62), and taking into account the estimate (9.63) and the relation $\zeta \sim \sigma(\zeta)$, we find that

$$\ln x_0 \gg e^{a^2\sigma^2/2} ,$$

Multiplying both sides of this inequality by x_0 and using (9.40), we have

$$ai_0R \gg x_0 \tag{9.65}$$

or

$$ai_0R \ln (ai_0R) \gg x_0 \ln x_0 .$$

Using (9.40) again, and dividing by ai_0R, we write the small-nonlinearity condition in the form

$$\ln (ai_0R) \gg e^{a^2\sigma^2/2} . \tag{9.66}$$

Thus, the method of small nonlinearity is applicable for large values of the parameter ai_0R and for noise levels which are not too large. In the case of a small correlation time τ_{cor}, corresponding to the sign $<$ in (9.64), we can replace (9.65) and (9.66) by less stringent conditions. Then, the region in which the method of small nonlinearity is applicable, which is essentially determined by the condition $\sigma(x) \ll x_0$, becomes somewhat larger.

3. The Method of the Fokker-Planck Equation

When the input process has a sufficiently small correlation time, i.e., one which is smaller than the other time constants of the system, we can analyze the process $x(t)$ by using stochastic methods involving the Fokker-Planck equation. We shall illustrate this method by applying it to the previously considered problem of the exponential detector, described by equation (9.36), which we write in the form

$$RC\dot{x} + x \ln x = ai_0 R\langle e^{a\xi}\rangle + ai_0 R\zeta . \tag{9.67}$$

When τ_{corr} is small, we can replace the random perturbation

$$\zeta = e^{a\xi} - \langle e^{a\xi}\rangle \tag{9.68}$$

in (9.67) by a Gaussian delta-correlated random process $\hat{\zeta}$, which has mean value zero (like ζ) and correlation function

$$\langle \hat{\zeta}\hat{\zeta}_\tau\rangle = K\delta(\tau) . \tag{9.69}$$

Here, the intensity coefficient K is chosen equal to the intensity coefficient of the original process (9.68):

$$K = \int_{-\infty}^{\infty} \langle \zeta\zeta_\tau\rangle \, d\tau . \tag{9.70}$$

After we have made this replacement of ζ by $\hat{\zeta}$, equation (9.67) describes a Markov process, and is equivalent to the Fokker-Planck equation

$$\dot{w} = \frac{1}{RC}\frac{\partial}{\partial x}[(x \ln x - m)\, w] + \frac{1}{2}\left(\frac{ai_0}{C}\right)^2 K \frac{\partial^2 w}{\partial x^2} , \tag{9.71}$$

where

$$m = ai_0 R\langle e^{a\xi}\rangle = ai_0 R e^{a^2\sigma^2/2} \tag{9.72}$$

[cf. (9.39)].

Finding the nonstationary solution of equation (9.71) is a complicated problem, but the one-dimensional stationary

probability density $w(x)$ can be found quite easily. Applying (4.49), we have

$$w(x) = N^{-1} \exp\left\{-\frac{RC}{(ai_0R)^2K}\left[x^2\left(\ln x - \frac{1}{2}\right) - 2mx\right]\right\}, \qquad (9.73)$$

where N is a suitable normalization constant. It is clear that the form of the distribution (9.73) is determined by only two parameters

$$m = ai_0Re^{a^2\sigma^2/2} \quad \text{and} \quad q = \frac{(ai_0R)^2}{2RC}K. \qquad (9.74)$$

The first parameter m determines the coordinate x_0 with maximum probability, and in fact x_0 is just the quantity satisfying equation (9.40), i.e., $x_0 \ln x_0 = m$. The second parameter q determines the amount of "scatter" due to the fluctuations. Thus, the coordinates x_1 and x_2, for which the probability density falls off to $e^{-1/2} \approx 0.61$ times its maximum value, can be found as the roots of the equation

$$f(x) - f(x_0) = q, \qquad (9.75)$$

where

$$f(x) = x^2(\ln x - \tfrac{1}{2}) - 2mx. \qquad (9.76)$$

From the distribution (9.73), we can easily find the probability density of the output voltage η, which is related to x by the zero-memory nonlinear transformation

$$\eta = \frac{1}{a}\ln x.$$

Using (9.73) and formula (1.15), we obtain

$$w(\eta) = \frac{a}{N}\exp\left\{-\frac{1}{2q}\left[e^{2a\eta}(a\eta - \tfrac{1}{2}) - 2me^{a\eta}\right] + a\eta\right\}. \qquad (9.77)$$

Because of the presence of the term $a\eta$ in the exponential, whose appearance is due to the derivative $dx/d\eta$, the value η_m corresponding to the maximum of the probability density (9.77) does not equal

$$\eta_0 = \frac{1}{a}\ln x_0.$$

In fact, η_m is now defined as the root of the equation

$$a\eta e^{2a\eta} - me^{a\eta} - q = 0. \tag{9.78}$$

Making a Taylor's series expansion of the function which is the argument of the exponential in (9.77) and takes its maximum value at $\eta = \eta_m$, we obtain

$$\begin{aligned} &-\frac{1}{2q}[e^{2a\eta}(a\eta - \tfrac{1}{2}) - 2me^{a\eta}] + a\eta \\ &= -\frac{1}{2q}[e^{2a\eta_m}(a\eta_m - \tfrac{1}{2}) - 2me^{a\eta_m}] + a\eta_m \\ &\quad -\frac{1}{2q}[e^{2a\eta_m}(2a\eta_m + 1) - me^{a\eta_m}]\, a^2(\eta - \eta_m)^2 + \ldots \end{aligned} \tag{9.79}$$

Retaining only these first few terms, we find that

$$w(\eta) \approx w(\eta_m) \exp\left\{-\frac{(\eta - \eta_m)^2}{2\sigma_{eq}^2}\right\} \tag{9.80}$$

in the region of maximum probability. Here,

$$\begin{aligned} \sigma_{eq}^2 &= \frac{q}{a^2}[e^{2a\eta_m}(2a\eta_m + 1) - me^{a\eta_m}]^{-1} \\ &= \frac{q}{a^2}\left[q\left(2 + \frac{1}{a\eta_m}\right) + me^{a\eta_m}\left(1 + \frac{1}{a\eta_m}\right)\right]^{-1} \end{aligned} \tag{9.81}$$

[cf. (9.78)] is the variance of an equivalent Gaussian distribution which is close to $w(\eta)$ in the region of maximum probability.

If the distribution $w(\eta)$ differs only slightly from a normal distribution, the quantity η_m is approximately equal to the mean value $\langle\eta\rangle$, and the equivalent variance σ_{eq}^2 is approximately equal to the variance $\mathbf{D}\eta$. For distributions which are not approximately normal, η_m and σ_{eq}^2 are no longer close to $\langle\eta\rangle$ and $\mathbf{D}\eta$, respectively, but they still serve to characterize the distribution $w(\eta)$. The equivalent variance σ_{eq}^2 characterizes the amount of "fluctuational scatter," in the sense that $\mathbf{D}\eta \sim \sigma_{eq}^2$. If we want to determine the actual mean value and variance in the case where $w(\eta)$ is not

approximately normal, we can calculate the appropriate integrals involving the weight (9.77), resorting to numerical integration if necessary.

For Gaussian noise $\xi(t)$, with

$$\langle \xi \rangle = 0 , \qquad \langle \xi \xi_\tau \rangle = \sigma^2 R(\tau) ,$$

we have

$$K = 2e^{a^2\sigma^2} \int_0^\infty [e^{a^2\sigma^2 R(\tau)} - 1]\, d\tau , \tag{9.82}$$

according to (9.54) and (9.70). For some specific forms of the correlation coefficient $R(\tau)$, we can immediately calculate this last integral. For example, if

$$R(\tau) = e^{-\gamma|\tau|} , \qquad \gamma = \frac{1}{\tau_{cor}} , \tag{9.83}$$

then, making the change of variable

$$z = a^2\sigma^2 e^{-\gamma\tau} ,$$

we find that

$$K = \frac{2}{\gamma} e^{a^2\sigma^2} \int_0^{a^2\sigma^2} \frac{e^z - 1}{z}\, dz = \frac{2}{\gamma} e^{a^2\sigma^2} [E^*(a^2\sigma^2) - 2 \ln a\sigma - C] . \tag{9.84}$$

Here, $E^*(z)$ is one of the exponential integrals,[2] and $C = 0.577...$ is Euler's constant. If the integral (9.82) cannot be calculated directly, the expansion

$$K = 2e^{a^2\sigma^2} \sum_{n=1}^{\infty} \frac{(a\sigma)^{2n}}{n!} \int_0^\infty R^n(\tau)\, d\tau \tag{9.85}$$

may turn out to be useful.

As an example, suppose that the parameters have the values

$$a\sigma = 3 , \qquad ai_0 R = 10 , \qquad \gamma RC = \frac{RC}{\tau_{cor}} = 100 .$$

[2] See Bateman Manuscript Project, *op. cit.*, p. 143.

Then

$$m = 898, \qquad q = 8.25 \times 10^6,$$

and, using (9.41), (9.78) and (9.81), we find that

$$a\eta_0 = 5.16, \qquad a\eta_m = 7.05, \qquad a\sigma_{\eta} = 0.655.$$

On the other hand, if $a\sigma$ and ai_0R have the same values, while the capacity is increased 10 times, so that $\gamma RC = 1000$, then

$$q = 8.25 \times 10^5, \qquad a\eta_m = 6.1, \qquad a\sigma_{\eta} = 0.543.$$

By comparing these results, we see that when the capacity C is increased, both the mean value of the voltage across the load and its fluctuational scatter decrease.

From the above quantitative results, pertaining to Gaussian noise with the specific correlation function (9.83) we can also conclude that the readings of a voltmeter which measures the mean value of the output voltage will be different for different values of τ_{cor}/RC, even if the variance σ^2 of the input noise remains the same. Therefore, it must be admitted that the use made of vacuum-tube voltmeters in ordinary engineering practice, to measure noise variances, is erroneous. Vacuum-tube voltmeters can be used only for comparative estimates of noise intensities, and only for noise whose spectral composition does not change. Thermoelectric devices should be used to measure actual mean square values of the noise fluctuations.

4. Detection of a Sine Wave Plus Noise with a Small Correlation Time

The method for calculating the one-dimensional distribution of the output of a nonlinear device with memory, presented in the preceding sections, is also applicable when the input contains a "sine wave" $E\cos\omega_0 t$, together with the random noise $\xi(t)$.

Then, for a detector with the exponential characteristic (9.33), the corresponding version of equation (9.36) is

$$RC\dot{x} + x \ln x = ai_0 R \exp\{aE \cos \omega t + a\xi\} . \tag{9.86}$$

We now consider two special cases:

1. Let the input noise process $\xi(t)$ be slowly varying compared to the rapidly oscillating signal $E \cos \omega_0 t$. We assume that the period $T_0 = 2\pi/\omega_0$ of these oscillations is much less than both the correlation time τ_{cor} of the noise and the filter time constant RC. Then equation (9.86) can be averaged over a period, which gives

$$RC\dot{x} + x \ln x = ai_0 R I_0(aE)\, e^{a\xi} . \tag{9.87}$$

This equation differs from (9.36) only by having the parameter $ai_0 R I_0(aE)$ instead of $ai_0 R$. Therefore, with the corresponding changes, everything said in Secs. 2 and 3 applies to this case.

2. The case where the process $\xi(t)$ is narrow-band is somewhat more distinctive. We assume that the frequency band of $\xi(t)$ lies near the frequency ω_0 of the sine wave. Then, the total input signal can be represented as an oscillation with resultant amplitude $B(t)$ and phase $\chi(t)$, i.e.,

$$E \cos \omega_0 t + \xi(t) = B(t) \cos[\omega_0 t + \psi(t)] . \tag{9.88}$$

We have already encountered such signals in Chap. 7 [cf. formula (7.80)]. If $\xi(t)$ is a Gaussian process, then the probability density of the amplitude B is of the form (7.88).

Substituting (9.88) into (9.86), we can use the fact that $B(t)$ and $\psi(t)$ are slowly varying compared to a sine wave of frequency ω_0. This allows us to average the right-hand side of (9.86) over a period $T_0 = 2\pi/\omega_0$, by regarding $B(t)$ and $\psi(t)$ as constants. The result is

$$RC\dot{x} + x \ln x = ai_0 R I_0[aB(t)] . \tag{9.89}$$

When the time constant RC greatly exceeds the time in which the amplitude B changes, i.e.,

$$\tau_{cor} \ll RC ,$$

where τ_{cor} is the correlation time of the noise, we can apply stochastic methods involving the Fokker-Planck equation. Thus, we regard the "fluctuational component"

$$\zeta(t) = I_0(aB) - \langle I_0(aB)\rangle \tag{9.90}$$

in (9.89) as being delta-correlated, by making the change

$$\langle \zeta\zeta_\tau\rangle \to K\delta(\tau)\,, \tag{9.91}$$

where

$$K = \int_{-\infty}^{\infty} \langle \zeta\zeta_\tau\rangle\, d\tau \tag{9.92}$$

is the intensity coefficient of the process (9.90). It is convenient to write

$$m = ai_0R\langle I_0(aB)\rangle\,, \tag{9.93}$$

since then (9.89) has the same form

$$RC\dot{x} + x\ln x = m + ai_0R\zeta \tag{9.94}$$

as in the preceding section, so that formulas (9.73) and (9.77) to (9.81) can be applied to the present case without any changes at all.

The parameters (9.92) and (9.93) of the process $\zeta(t)$ involve the averages $\langle I_0(aB)\rangle$ and $\langle I_0(aB)I_0(aB_\tau)\rangle$. To calculate these quantities, it is convenient to average the random function

$$e^{ax(t)} = \exp\{E\cos\omega_0 t + a\xi(t)\}$$

before carrying out the time average over a period. Thus, if $\xi(t)$ is a Gaussian random process with mean value zero and correlation function

$$\sigma^2R(\tau) = \sigma^2 r(\tau)\cos\omega_0\tau\,,$$

we have

$$\langle e^{ax}\rangle = \exp\{aE\cos\omega_0 t + \tfrac{1}{2}a^2\sigma^2\}\,, \tag{9.95}$$

$$\langle e^{ax+ax_\tau}\rangle = \exp\{aE\cos\omega_0 t + aE\cos\omega_0(t+\tau) + a^2\sigma^2[1 + r(\tau)\cos\omega_0\tau]\}\,.$$

Averaging the first of these expressions with respect to t over a period, we obtain

$$\langle I_0(aB)\rangle = \frac{\omega_0}{2\pi}\int_t^{t+(2\pi/\omega_0)} \langle e^{ax(t')}\rangle\, dt' = e^{a^2\sigma^2/2}\, I_0(aE).$$

Averaging the second of the equations (9.95) with respect to t over a period is equivalent to averaging the expression (8.121) with respect to a random initial phase. Setting $i\Omega = i\Omega_\tau = a$ in (8.121) and (8.135), we find that

$$\langle e^{ax+ax_\tau}\rangle = e^{a^2\sigma^2}\sum_{k=0}^{\infty}\epsilon_k I_k^2(aE)\, e^{a^2\sigma^2 r\cos\omega_0\tau}\cos k\omega_0\tau\,. \tag{9.96}$$

We now average (9.96) with respect to τ. According to (8.136), this gives

$$e^{a^2\sigma^2}\sum_{k=0}^{\infty}\epsilon_k I_k^2(aE)\, I_k(a^2\sigma^2 r)\,.$$

Therefore, we have

$$\langle I_0(aB)\, I_0(aB_\tau)\rangle = e^{a^2\sigma^2}\sum_{k=0}^{\infty}\epsilon_k I_k^2(aE)\, I_k[a^2\sigma^2 r(\tau)]\,, \tag{9.97}$$

$$\langle\zeta\zeta_\tau\rangle = e^{a^2\sigma^2} I_0^2(aE)\{I_0[a^2\sigma^2 r(\tau)] - 1\} + 2e^{a^2\sigma^2}\sum_{k=1}^{\infty} I_k^2(aE)\, I_k[a^2\sigma^2 r(\tau)]\,. \tag{9.98}$$

To calculate the intensity coefficient K, we substitute (9.98) into (9.92). It may be convenient to use a series expansion of the Bessel function, when evaluating the integral with respect to τ. Suppose the correlation coefficient is of the form

$$r(\tau) = e^{-|\gamma|\tau}\,.$$

Then we find it useful to introduce the new variable of integration

$$z = a^2\sigma^2 e^{-\gamma\tau}\,.$$

In the special case $E = 0$, according to (9.98) and (9.92), we have

$$K = \frac{2}{\gamma}e^{a^2\sigma^2}\int_0^{a^2\sigma^2}\frac{I_0(z) - 1}{z}\,dz = \frac{1}{\gamma}e^{a^2\sigma^2}\sum_{m=1}^{\infty}\frac{1}{m(m!)^2}\left(\frac{a^2\sigma^2}{2}\right)^{2m}. \tag{9.99}$$

When $a^2\sigma^2 \gg 1$, it is convenient to use the asymptotic representation

$$\int \frac{I_0(z)}{z}\,dz \sim \frac{I_1(z)}{z} + 2\frac{I_2(z)}{z^2} + \ldots + n!\,2^n \frac{I_{n+1}(z)}{z^{n+1}} + \ldots \tag{9.100}$$

It is interesting to compare the expression (9.99) with formula (9.84), pertaining to the case of slowly varying input fluctuations. For the value $a\sigma = 2$, the intensity coefficient (9.84) is

$$K \approx \frac{2070}{\gamma},$$

while, for "quasi-harmonic" input fluctuations, (9.99) gives

$$K = \frac{1}{\gamma} e^4(4 + 2 + 0.59 + 0.11 + 0.03 + \ldots) \approx \frac{366}{\gamma}.$$

5. The Case of Intermediate Correlation Times

The most difficult case is where the correlation time of the input process $\xi(t)$ in the equation

$$\dot{\eta} + \frac{1}{RC}\eta = \frac{1}{C}F(\xi - \eta) \tag{9.101}$$

is comparable with the filter time constant RC. Then, we can use neither the quasi-static model, nor the method of the Fokker-Planck equation in the form given above. However, in this case, we can still use stochastic methods involving the Fokker-Planck equation, if we consider two-dimensional Markov processes.

Thus, let $\xi(t)$ be a Gaussian process with mean value zero and correlation function

$$\langle \xi\xi_\tau \rangle = \sigma^2 e^{-\gamma|\tau|}. \tag{9.102}$$

As we know from Chap. 4, Sec. 10, such a process is Markovian and satisfies the differential equation

$$\dot{\xi} + \gamma\xi = \zeta(t), \tag{9.103}$$

where $\zeta(t)$ is a Gaussian delta-correlated process such that

$$\langle \zeta \rangle = 0, \qquad \langle \zeta \zeta_\tau \rangle = 2\gamma\sigma^2\delta(\tau) = K_0\delta(\tau).$$

Equations (9.101) and (9.103) together specify a two-dimensional Markov process. Writing $v = \xi - \eta$, we can replace these equations by the new system

$$\begin{aligned} \dot{v} &= -\frac{v}{RC} - \frac{1}{C}F(v) - \left(\gamma - \frac{1}{RC}\right)\xi + \zeta, \\ \dot{\xi} &= -\gamma\xi + \zeta, \end{aligned} \tag{9.104}$$

with the equivalent Fokker-Planck equation

$$\begin{aligned} \dot{w}(v, \xi) = {} & \frac{\partial}{\partial v}\left\{\left[\frac{v}{RC} + \frac{1}{C}F(v)\right]w\right\} + \left(\gamma - \frac{1}{RC}\right)\xi\frac{\partial w}{\partial v} \\ & + \gamma\frac{\partial}{\partial \xi}(\xi w) + \gamma\sigma^2\left(\frac{\partial^2 w}{\partial v^2} + 2\frac{\partial^2 w}{\partial v \partial \xi} + \frac{\partial^2 w}{\partial \xi^2}\right). \end{aligned} \tag{9.105}$$

Using this equation to find the probability density $w(v, \xi)$ or $w(v, \xi; v_\tau, \xi_\tau)$, we can then integrate with respect to ξ and ξ_τ to obtain

$$w(\eta) = \int w(\xi - \eta, \xi)\, d\xi,$$

$$w(\eta, \eta_\tau) = \int w(\xi - \eta, \xi; \xi_\tau - \eta_\tau, \xi_\tau)\, d\xi\, d\xi_\tau.$$

Unfortunately, the complete solution of equation (9.103) is a difficult problem, and hence we shall confine ourselves to special cases.

5.1. First, we consider the case where the correlation time is exactly equal to the filter time constant, so that

$$\gamma = \frac{1}{RC}. \tag{9.106}$$

Then the voltage across the diode satisfies the equation

$$\dot{v} = -\frac{v}{RC} - \frac{1}{C}F(v) + \zeta(t), \tag{9.107}$$

and since it does not contain ξ, this equation can be solved independently. From the process $v(t)$, we can obtain the output voltage by using (9.101). The result is

$$\eta(t) = \frac{1}{C}\int_{-\tau}^{t} \exp\left\{-\frac{t-t'}{RC}\right\} F[v(t')]\, dt' \,. \tag{9.108}$$

since

$$\dot{\eta} + \frac{1}{RC}\eta = \frac{1}{C}F \,.$$

The average of the stationary value of η equals

$$\langle \eta \rangle = R\langle F(v) \rangle \,, \tag{9.109}$$

and to find the output correlation function, it is sufficient to know

$$\langle F(v)\, F(v_\tau) \rangle \,.$$

The Fokker-Planck equation corresponding to (9.107) is

$$\dot{w}(v) = \frac{1}{RC}\frac{\partial}{\partial v}\{[v + RF(v)]\, w\} + \gamma\sigma^2 \frac{\partial^2 w}{\partial v^2} \,, \tag{9.110}$$

which has the stationary solution

$$w(v) = \frac{1}{N}\exp\left\{-\frac{v^2}{2\sigma^2} - \frac{R}{\sigma^2}\int_{-\tau}^{v} F(z)\, dz\right\} \,. \tag{9.111}$$

To find $\langle F(v) \rangle$, we have to evaluate the integrals

$$\langle F(v) \rangle = \frac{1}{N}\int_{-\infty}^{\infty} \exp\left\{-\frac{v^2}{2\sigma^2} - \frac{R}{\sigma^2}\int_{-\tau}^{v} F(z)\, dz\right\} F(v)\, dv \tag{9.112}$$

and

$$N = \int_{-\infty}^{\infty} \exp\left\{-\frac{v^2}{2\sigma^2} - \frac{R}{\sigma^2}\int_{-\tau}^{v} F(z)\, dz\right\} dv \,.$$

If the current-voltage characteristic of the diode is piecewise linear, of the form

$$F(v) = \begin{cases} Sv & \text{for } v \geq 0 \,, \\ 0 & \text{for } v < 0 \,, \end{cases} \tag{9.113}$$

these formulas lead to the results

$$\langle F(v)\rangle = \frac{S}{N}\frac{\sigma^2}{1+RS}, \qquad N = \sqrt{\frac{\pi}{2}}\left[1+(1+RS)^{-1/2}\right], \tag{9.114}$$

and hence, according to (9.109),

$$\langle \eta\rangle = \sqrt{\frac{2}{\pi}}\frac{RS\sigma}{1+RS+\sqrt{1+RS}}. \tag{9.115}$$

We get different expressions for $\langle F(v)\rangle$, N and $\langle \eta\rangle$ if the characteristic $F(v)$ has a different form. In what follows, we shall not specify the precise form of the function $F(v)$, but we shall assume that it vanishes for $v<0$, and grows rapidly for $v>0$, in fact, much more rapidly than v/R. This second condition is satisfied in most practical cases, where the diode has a small internal resistance $R_i \ll R$ in its conducting state, so that the ratio of the output voltage of the detector to the input voltage is close to unity. Under these conditions, the probability

$$w_n = \int_{-\infty}^{0} w(v)\,dv = \sqrt{\frac{\pi}{2}}\frac{\sigma}{N} \tag{9.116}$$

that the diode is in the nonconducting state greatly exceeds the probability

$$w_c = \int_0^{\infty} w(v)\,dv \approx \frac{1}{N}\int_0^{\infty}\exp\left\{-\frac{R}{\sigma^2}\int_0^{v}F(z)\,dz\right\}dv \tag{9.117}$$

that it is in the conducting state. Moreover, the inequality

$$w_c \ll w_n \tag{9.118}$$

shows that most of the time the diode is in the nonconducting state. In fact, it is clear that

$$w_n \approx 1, \qquad N \approx \sigma\sqrt{\frac{\pi}{2}}. \tag{9.119}$$

To calculate the average (9.112), we can neglect the term $v^2/2$ in the exponential compared to $R\int_0^v F(z)\,dz$, obtaining

$$\langle F(v)\rangle = \sqrt{\frac{2}{\pi}}\frac{1}{\sigma}\int_0^\infty \exp\left\{-\frac{R}{\sigma^2}\int_{-v}^{v} F(z)\,dz\right\} F(v)\,dv \approx \sqrt{\frac{2}{\pi}}\frac{\sigma}{R}, \tag{9.120}$$

so that

$$\langle \eta\rangle = \sqrt{\frac{2}{\pi}}\,\sigma.$$

Turning to the calculation of the correlation function, we fix the initial position $v(0) = v_0$ of the "representative point" (see p. 63). Suppose that $v_0 < 0$. Then, according to (9.107), we have

$$\dot v + \gamma v = \zeta(t), \tag{9.121}$$

as long as v is negative, since the function $F(v)$ has no effect in the region where v is negative. Using (9.121), we can find the average displacement and the "fluctuational scatter" of the representative point:

$$\begin{aligned} \langle v\rangle &= v_0 e^{-\gamma t}, \\ \mathbf{D}v &= \sigma^2(1 - e^{-2\gamma t}). \end{aligned} \tag{9.122}$$

Therefore, $v(t)$ is described by the probability density

$$w(v, t \mid v_0) = \frac{1}{\sqrt{2\pi}}\,\frac{1}{\sigma\sqrt{1 - e^{-2\gamma t}}}\exp\left\{-\frac{1}{2\sigma^2}\,\frac{(v - v_0 e^{-\gamma t})^2}{1 - e^{-2\gamma t}}\right\}, \tag{9.123}$$

which is the solution of the Fokker-Planck equation

$$\dot w = \gamma\frac{\partial(vw)}{\partial v} + \gamma\sigma^2\frac{\partial^2 w}{\partial v^2}. \tag{9.124}$$

When the representative point reaches the region where $v > 0$, the function $F(v)$ comes into play, and reflects the representative point back into the region where $v < 0$. In fact, there remains

only a very small probability $w_c \ll 1$ of penetration into the region where $v > 0$. Therefore, we can assume that the probability current (see p. 62) through the boundary $v = 0$ vanishes, i.e.,

$$-\gamma v w - \gamma\sigma^2 \frac{\partial w}{\partial v} = 0 \quad \text{for} \quad v = 0\,,$$

which leads to the boundary condition

$$\frac{\partial w\,(0)}{\partial v} = 0\,. \tag{9.125}$$

This condition is not satisfied by the function (9.123); the solution of (9.124) which does satisfy (9.125) has the form

$$w(v, t \mid v_0) = \frac{1}{\sqrt{2\pi}\sigma} \frac{1}{\sqrt{1 - e^{-2\gamma t}}} \times \left[\exp\left\{-\frac{1}{2\sigma^2}\frac{(v - v_0 e^{-\gamma t})^2}{1 - e^{-2\gamma t}}\right\} + \exp\left\{-\frac{1}{2\sigma^2}\frac{(v + v_0 e^{-\gamma t})^2}{1 - e^{-2\gamma t}}\right\}\right] \tag{9.126}$$

$$(v < 0\,,\ v_0 < 0)\,.$$

Then at the origin we have

$$w(0, t \mid v_0) = \frac{1}{\sqrt{2\pi}\sigma} \frac{2}{\sqrt{1 - e^{-2\gamma t}}} \exp\left\{-\frac{1}{2\sigma^2}\frac{v_0^2 e^{-2\gamma t}}{1 - e^{-2\gamma t}}\right\}, \tag{9.127}$$

and the function is extended in a continuous fashion into the region $v > 0$. In this region, the distribution is practically the same as the stationary distribution

$$w(v, t \mid v_0) = \frac{1}{\sqrt{2\pi}\sigma} \frac{2}{\sqrt{1 - e^{-2\gamma t}}} \exp\left\{-\frac{1}{2\sigma^2}\frac{v_0^2 e^{-2\gamma t}}{1 - e^{-2\gamma t}}\right\} \times \exp\left\{-\frac{v^2}{2\sigma^2} - \frac{R}{\sigma^2}\int_0^v F(z)\,dz\right\} \tag{9.128}$$

$$(v > 0\,,\ v_0 < 0)\,,$$

which has the same value at the origin as (9.126).

If the initial point approaches the boundary, i.e., $v_0 \to 0$, the function (9.126), (9.128) goes into

$$w(v, t \mid 0) = \begin{cases} \sqrt{\frac{2}{\pi}} \frac{1}{\sigma \sqrt{1 - e^{-2\gamma t}}} \exp\left\{ -\frac{1}{2\sigma^2} \frac{v^2}{1 - e^{-2\gamma t}} \right\} & \text{for } v < 0, \\ \sqrt{\frac{2}{\pi}} \frac{1}{\sigma \sqrt{1 - e^{-2\gamma t}}} \exp\left\{ -\frac{v^2}{2\sigma^2} - \frac{R}{\sigma^2} \int_0^v F(z)\, dz \right\} & \text{for } v > 0. \end{cases} \tag{9.129}$$

When the initial value v_0 lies in the region where v is positive (so that $v_0 > 0$), then, due to the influence of $F(v)$, the representative point is rapidly "pushed out" into the region where v is negative. After this occurs, the distribution $w(v \mid v_0)$ is close to the distribution (9.129) just found. In fact, the two distributions differ only during a time of order $(\gamma RF')^{-1} = R_i C$, which is negligibly small compared to $1/\gamma = RC$, because of the assumption made above.

Once we know the one-dimensional distribution (9.111) and the transition probability $w(v, t \mid v_0)$, we can write down the two-dimensional probability density

$$w(v, v_\tau) = w(v)\, w(v_\tau, \tau \mid v)$$

and the second moment

$$\langle F(v) F(v_\tau) \rangle = \int_0^\infty \int_0^\infty F(v) F(v_\tau)\, w(v)\, w(v_\tau, \tau \mid v)\, dv\, dv_\tau . \tag{9.130}$$

Taking $w(v_\tau, \tau \mid v)$ in (9.130) to be (9.129), we find that

$$\begin{aligned} \langle F(v) F(v_\tau) \rangle &= \frac{1}{N} \sqrt{\frac{2}{\pi}} \frac{1}{\sigma \sqrt{1 - e^{-2\gamma\tau}}} \\ &\quad \times \left[\int_0^\infty F(v) \exp\left\{ -\frac{v^2}{2\sigma^2} - \frac{R}{\sigma^2} \int_0^v F(z)\, dz \right\} dv \right]^2 \qquad (9.131) \\ &= N \sqrt{\frac{2}{\pi}} \frac{\langle F(v) \rangle^2}{\sigma \sqrt{1 - e^{-2\gamma\tau}}} \qquad (\tau > 0) . \end{aligned}$$

Because of (9.119) and (9.120), this gives

$$k_F(\tau) = \langle F(v)\,F(v_\tau)\rangle - \langle F(v)\rangle^2 = \frac{2}{\pi}\frac{\sigma^2}{R^2}\left[\frac{1}{\sqrt{1-e^{-2\gamma|\tau|}}} - 1\right]. \quad (9.132)$$

For small $\tau < R_iC$, the expression (9.132) is not valid, and for $\tau = 0$, we can use the one-dimensional distribution (9.111), which leads to

$$k_F(0) = \mathbf{D}F(v) = \sqrt{\frac{2}{\pi}}\frac{1}{\sigma}\int_0^\infty \exp\left\{-\frac{R}{\sigma^2}\int_0^v F(z)\,dz\right\} F^2(v)\,dv. \quad (9.133)$$

To find the spectral density $S[F(v) - \langle F\rangle; \omega]$ for frequencies ω of order γ and smaller, we can take the Fourier transform of (9.132). The region of small values of τ, where (9.132) is not valid, will have an important effect only for high frequencies $\omega \sim 1/R_iC$. The Fourier transform of (9.131) can be found by using the obvious relation

$$\int_{-\infty}^{\infty} e^{i\omega\tau}\left[\frac{1}{\sqrt{1-e^{-2\gamma|\tau|}}} - 1\right]d\tau = 2\,\mathrm{Re}\int_0^\infty e^{-p\tau}\left[\frac{1}{\sqrt{1-e^{-2\gamma\tau}}} - 1\right]d\tau,$$

where $p = i\omega$, and the formula[3]

$$\int_0^\infty e^{-p\tau}\frac{d\tau}{\sqrt{1-e^{-2\gamma\tau}}} = \frac{1}{2\gamma}\int_0^\infty e^{-pt/2\gamma}\frac{dt}{\sqrt{1-e^{-t}}} = \frac{1}{2\gamma}B\left(\frac{p}{2\gamma}, \frac{1}{2}\right), \quad (9.134)$$

where $B(x, y)$ is the beta function. Comparing the last two equations, we see that

$$\int_{-\infty}^{\infty} e^{i\omega\tau}\left[\frac{1}{\sqrt{1-e^{-2\gamma|\tau|}}} - 1\right]d\tau = \frac{1}{\gamma}\,\mathrm{Re}\left[B\left(\frac{i\omega}{2\gamma}, \frac{1}{2}\right) - \frac{2\gamma}{i\omega}\right],$$

and hence, according to (2.11) and (9.132),

$$S[F(v) - \langle F\rangle; \omega] = \frac{4}{\pi}\frac{\sigma^2 C}{R}\,\mathrm{Re}\,B\left(\frac{i\omega RC}{2}, \frac{1}{2}\right). \quad (9.135)$$

[3] See I. M. Ryshik and I. S. Gradstein, *op. cit.*, formula (5.218.3), p. 258.

To determine the spectral density numerically, we can use tables of the gamma function, recalling that

$$B\left(iz, \tfrac{1}{2}\right) = \sqrt{\pi}\,\frac{\Gamma(iz)}{\Gamma\left(iz + \frac{1}{2}\right)},$$

or else we can use the expansion

$$B\left(iz, \tfrac{1}{2}\right) = \sum_{k=0}^{\infty} \frac{(2k-1)!}{2^{2k-1}k!(k-1)!}\,\frac{1}{iz+k}.$$

To improve the convergence of the last formula, we transform it into[4]

$$B\left(iz, \tfrac{1}{2}\right) = \frac{1}{iz} + 2\ln 2 - iz\sum_{k=1}^{\infty} \frac{(2k-1)!}{2^{2k-1}k!(k-1)!}\,\frac{1}{k(iz+k)},$$

which implies that

$$\operatorname{Re} B\left(iz, \tfrac{1}{2}\right) = 2\ln 2 - z^2\sum_{k=1}^{\infty} \frac{(2k-1)!}{2^{2k-1}k!(k-1)!}\,\frac{1}{k(k^2+z^2)}. \quad (9.136)$$

Substituting (9.136) into (9.135), we have

$$S[F(v) - \langle F\rangle\,;\omega] = \frac{4}{\pi}\,\frac{\sigma^2 C}{R}\left\{2\ln 2 - z^2\left[\frac{1}{2}\,\frac{1}{1+z^2} + \frac{1\cdot 3}{2\cdot 4}\,\frac{1}{2(4+z^2)}\right.\right.$$

$$\left.\left. + \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\,\frac{1}{3(9+z^2)} + \cdots\right]\right\},$$

[4] Use the fact that

$$2\ln 2 = \sum_{k=1}^{\infty} \frac{(2k-1)!}{2^{2k-1}(k!)^2},$$

which follows from

$$2\ln 2 = \int_0^1 \frac{(1-x)^{-1/2} - 1}{x}\,dx$$

if the integrand is expanded in a power series and integrated term by term.

where $z = \omega/2\gamma$. This function, which gives the spectral density of the fluctuations of the current $I = F(v)$, is plotted in Figure 15.

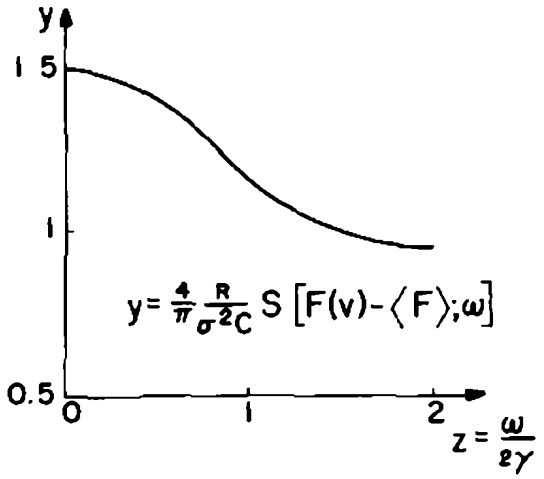

Fig. 15. The spectral density of the fluctuations of the current $I = F(v)$.

According to (9.101), the output voltage η is obtained from $F(v)$ by a linear transformation. It follows from the usual rules and (9.135) that the spectral density of the fluctuations of the output voltage is

$$S[\eta - \langle\eta\rangle\,;\,\omega] = \left|i\omega C + \frac{1}{R}\right|^{-2} S[F(v) - \langle F\rangle\,;\,\omega] = \frac{4}{\pi}\sigma^2 \frac{RC}{\omega^2 R^2 C^2 + 1} \operatorname{Re} B\left(\frac{i\omega RC}{2}, \frac{1}{2}\right). \tag{9.137}$$

Finally, according to (2.15), to find the variance $\mathbf{D}\eta$, we have to integrate (9.137) with respect to the frequency[5]:

$$\mathbf{D}\eta = \frac{2}{\pi^2}\sigma^2 \int_0^\infty \operatorname{Re} B\left(\frac{ix}{2}, \frac{1}{2}\right)\frac{dx}{x^2+1} = \left(1 - \frac{2}{\pi}\right)\sigma^2 \approx 0.363\,\sigma^2\,. \tag{9.138}$$

[5] To evaluate the integral in (9.138), we use (9.136) and the easily verified formula

$$\int_0^\infty \frac{x^2 dx}{(x^2+1)(x^2+4k^2)} = \frac{\pi}{2}\frac{1}{2k+1},$$

5.2. When the noise correlation time $1/\gamma$ is different from RC, we have to solve the two-dimensionel Fokker-Planck equation (9.105). This can be done by using methods similar to those given in Chap. 4, Sec. 11. Thus, to find the stationary distribution satisfying the equation

$$\gamma\sigma^2 \frac{\partial^2 w}{\partial v^2} + \frac{\partial}{\partial v}\left(\left[\frac{v}{RC} + \frac{1}{C} F(v)\right] w\right) + \gamma\sigma^2 \frac{\partial^2 w}{\partial \xi^2} + \gamma \frac{\partial}{\partial \xi}(\xi w) + \epsilon\xi \frac{\partial w}{\partial v} + 2\gamma\sigma^2 \frac{\partial^2 w}{\partial v \partial \xi} = 0, \tag{9.139}$$

where

$$\epsilon = \gamma - \frac{1}{RC},$$

we look for a solution of the form

$$w(v, \xi) = \sum_{m,n=0}^{\infty} T_{mn} V_m(v) X_n(\xi). \tag{9.140}$$

Here, the functions

$$X_n(\xi) = \frac{1}{\sigma\sqrt{n!}} H_n\left(\frac{\xi}{\sigma}\right) \tag{9.141}$$

are the eigenfunctions of equation (4.68), which satisfy the relations

$$X_n' = \frac{\sqrt{n+1}}{\sigma} X_{n+1},$$
$$\xi X_n = -\sigma[\sqrt{n}\, X_{n-1} + \sqrt{n+1}\, X_{n+1}]. \tag{9.142}$$

---

obtaining

$$\int_0^\infty \operatorname{Re} B\left(\frac{ix}{2}, \frac{1}{2}\right) \frac{dx}{x^2+1} = \pi\left[\ln 2 - \sum_{k=1}^{\infty} \frac{(2k-1)!}{2^{2k-1}k!(k-1)!} \frac{1}{2k(2k+1)}\right].$$

According to I. M. Ryshik and I. S. Gradstein, *op. cit.*, formulas (1.641.1) and (1.642.2), p. 48,

$$\sum_{k=1}^{\infty} \frac{(2k-1)!}{2^{2k-1}k!(k-1)!} \frac{x^{2k+1}}{2k(2k+1)} = x \ln\left(\frac{2}{x} - \operatorname{arc\,cosh}\frac{1}{x}\right) - \arcsin x + x$$

for $x \leqslant 1$. If we set $x = 1$, (9.138) now follows from the last two formulas.

The eigenfunctions $V_m(v)$ satisfy the equation

$$\gamma\sigma^2 \frac{\partial^2 V_m}{\partial v^2} + \frac{\partial}{\partial v}\left(\left[\frac{v}{RC} + \frac{1}{C} F(v)\right] V_m\right) + \lambda_m V_m = 0. \tag{9.143}$$

Substituting (9.140) into (9.139), and taking into account (9.142) and (9.143), we obtain

$$\begin{aligned} &- \sum_{m,n} \lambda_m T_{mn} V_m X_n - \gamma \sum_{m,n} n T_{mn} V_m X_n \\ &- \epsilon\sigma \sum_{m,n} T_{mn} V_m' [\sqrt{n}\, X_{n-1} + \sqrt{n+1}\, X_{n+1}] \\ &+ 2\gamma\sigma \sum_{m,n} T_{mn} V_m' \sqrt{n+1}\, X_{n+1} = 0. \end{aligned} \tag{9.144}$$

Next, we expand the derivatives V_k with respect to the eigenfunctions V_m, by writing

$$V_k' = \sum_m V_m a_{mk},$$

where

$$a_{mk} = \int V_m V_k' \frac{dv}{V_0}. \tag{9.145}$$

Then, from (9.144) we obtain the equations

$$\begin{aligned} (\lambda_m + \gamma n)\, T_{mn} = (2\gamma - \epsilon)\, \sigma \sqrt{n} \sum_k a_{mk} T_{k,n-1} \\ - \epsilon\sigma \sqrt{n+1} \sum_k a_{mk} T_{k,n+1}, \end{aligned} \tag{9.146}$$

which can be solved by various methods. Having specified a given accuracy, we need only consider a certain number of expansion coefficients T_{mn} of low order, setting the higher-order coefficients equal to zero. Then, (9.146) gives a finite system of linear equations, from which we can determine the T_{mn}. Another way of solving the problem is to use the method of successive approximations, regarding ϵ as a small parameter, and obtaining the solution in the form of an expansion in powers of this parameter.

★ Combining the expressions obtained from (9.146) for $n = 0$, $n = 1$ and $n = 2$, we find that

$$T_{m0} = -\epsilon(2\gamma - \epsilon)\sigma^2 \sum_k b_{mk} T_{k0} + \epsilon^2(2\gamma - \epsilon)^2 \sigma^4 \sum_{k \neq 0, l} b_{mk} \frac{2\lambda_k}{\lambda_k + 2\gamma} b_{kl} T_{l0} + O(\epsilon^3) \qquad (m \neq 0), \tag{9.147}$$

where we have written

$$b_{mk} = \frac{1}{\lambda_m} \sum_j \frac{a_{mj} a_{jk}}{\lambda_j + \gamma}.$$

By repeatedly substituting the whole right-hand side of (9.147) for T_{k0}, T_{l0}, ..., we can express the coefficients T_{m0} in terms of T_{00}:

$$T_{m0} = -\epsilon(2\gamma - \epsilon)\sigma^2 b_{m0} T_{00} + \epsilon^2(2\gamma - \epsilon)^2 \sigma^4 \sum_{k \neq 0} b_{mk} \frac{3\lambda_k + 2\gamma}{\lambda_k + 2\gamma} b_{k0} T_{00} + O(\epsilon^3) \qquad (m \neq 0). \tag{9.148}$$

To determine the other coefficients T_{mk}, we again turn to the equations (9.146), applying the method of successive approximations. Letting $T_{mk}^{(0)}$ denote the zeroth approximation, i.e.,

$$T_{mk} = T_{mk}^{(0)} + O(\epsilon),$$

we find from (9.146) that

$$\begin{aligned} T_{m1}^{(0)} &= (2\gamma - \epsilon)\sigma \frac{a_{m0}}{\lambda_m + \gamma} T_{00}, \\ T_{m2}^{(0)} &= (2\gamma - \epsilon)\sigma \frac{\sqrt{2}}{\lambda_m + 2\gamma} \sum_k a_{mk} T_{k1}^{(0)} \\ &= (2\gamma - \epsilon)^2 \sigma^2 \frac{\sqrt{2}}{\lambda_m + 2\gamma} \sum_k \frac{a_{mk} a_{k0}}{\lambda_k + \gamma} T_{00}, \end{aligned} \tag{9.149}$$

The zeroth approximation is then used to find the first approximation to $T_{mk}^{(1)}$, which differs from T_{mk} by terms of order ϵ^2:

$$T_{m1}^{(1)} = T_{m1}^{(0)} + \frac{(2\gamma - \epsilon)\sigma}{\lambda_m + \gamma} \sum_{k \neq 0} a_{mk} T_{k0} - \epsilon\sigma \frac{\sqrt{2}}{\lambda_m + \gamma} \sum_k a_{mk} T_{k2}^{(0)},$$

$$T_{m2}^{(1)} = (2\gamma - \epsilon)\sigma \frac{\sqrt{2}}{\lambda_m + 2\gamma} \sum_k a_{mk} T_{k1}^{(1)} - \epsilon\sigma \frac{\sqrt{3}}{\lambda_m + 2\gamma} \sum_k a_{mk} T_{k3}^{(0)},$$

$$\cdots \tag{9.150}$$

This method allows us to find T_{mk} to within terms of order ϵ^μ, where μ is any integer. Naturally, the larger μ, the more formidable the calculations become.★

All the coefficients T_{mn} in the expansion (9.140) turn out to be proportional to the first coefficient T_{00}, which is determined by the normalization coefficient. In fact, if we integrate (9.140) with respect to v and ξ, all the terms except the first vanish, because of (4.54), and hence

$$\int w(v, \xi)\, dv\, d\xi = T_{00} = 1. \tag{9.151}$$

To simplify the calculations, it is convenient to assume (as on pp.274-277) that the function $F(v)$ vanishes for $v < 0$ and grows rapidly for $v > 0$. More precisely, let $R_i = (\partial F/\partial v)^{-1}$, the internal resistance of the diode in its conducting state, satisfy the conditions

$$R_i \ll R, \qquad R_i C \ll \frac{1}{\gamma}. \tag{9.152}$$

Then, in considering the region $v > 0$, we can neglect the terms

$$\frac{v}{RC}, \qquad \left(\gamma - \frac{1}{RC}\right)\xi$$

in the first of the equations (9.104), in comparison with $F(v)/C$; instead, we can use the equation

$$\dot{v} = -\frac{1}{C}F(v) + \zeta. \tag{9.153}$$

The corresponding Fokker-Planck equation has the stationary solution

$$w(v) = \text{const} \cdot \exp\left\{ -\frac{1}{\gamma\sigma^2 C} \int_0^v F(z)\, dz \right\}, \tag{9.154}$$

which is "established" after a time exceeding the small time constant $R_i C$.

For the region $v < 0$, we can replace the function $F(v)$ in equations (9.104) and (9.105), by the function

$$F_1(v) = \begin{cases} 0 & \text{for} \quad v < 0, \\ \infty & \text{for} \quad v > 0. \end{cases} \tag{9.155}$$

The infinite value of $F_1(v)$ for $v > 0$ means that the representative point is immediately reflected back when it reaches the boundary $v = 0$. Actually, the reflection takes a time of order $R_i C$, which is considerably less than the time constants $1/RC$ and $1/\gamma$ which characterize the process for $v < 0$.

After replacing $F(v)$ by (9.155), equation (9.143) has the eigenfunctions

$$V_m(v) = \frac{2}{\sqrt{(2m)!}} \frac{1}{\sigma\sqrt{\gamma RC}} F^{(2m+1)}\left(\frac{v}{\sigma\sqrt{\gamma RC}}\right) \quad (v < 0) \tag{9.156}$$

[see formula (4.70)], and the eigenvalues

$$\lambda_m = \frac{2m}{RC}.$$

According to (9.145), in the present case, using the formula[6]

$$\int_0^\infty H_{2m}(z)\, H_{2k+1}(z)\, e^{-z^2/2}\, dz = (-1)^{m-k+1} \frac{(2m-1)!\,(2k-1)!}{2^{m+k-1}(m-1)!\,k!\,(2m-2k-1)},$$

[6] Cf. Bateman Manuscript Project, *op. cit.*, formula (9), p. 117; formula (21), p. 122; formula (7), p. 3.

we have

$$a_{mk} = \frac{(-1)^{m+k+1}}{\sigma\sqrt{2\pi\gamma RC}} \frac{2}{\sqrt{(2m)!\,(2k)!}} \frac{(2m-1)!\,(2k+1)!}{2^{m+k-1}(m-1)!\,k!\,(2m-2k-1)}. \tag{9.157}$$

If we use (9.146) to calculate T_{mn} to some desired accuracy, then, according to (9.140), we obtain the joint stationary probability density of the input and output signals, i.e.,

$$w_{\eta\xi}(\eta, \xi) = [w(v, \xi)]_{v=\xi-\eta} = \sum_{m,n} T_{mn} V_m(\xi - \eta)\, X_n(\xi), \tag{9.158}$$

where the functions V_m and X_n are given by formulas (9.156) and (9.141). By using expansions of this kind, we can also investigate nonstationary transient processes, and two-dimensional probability densities corresponding to different instants of time.

Bibliography

Bharucha-Reid, A. T., *Elements of the Theory of Markov Processes and Their Applications*, McGraw-Hill Book Co., Inc., New York (1960).

Cramér, H., *Mathematical Methods of Statistics*, Princeton University Press, Princeton, N. J. (1946).

Davenport, W. B., Jr., and W. L. Root, *An Introduction to the Theory of Random Signals and Noise*, McGraw-Hill Book Co., Inc., New York (1958).

Deutsch, R., *Nonlinear Transformations of Random Processes*, Prentice-Hall, Inc., Englewood Cliffs, N. J. (1962).

Fano, R. M., *Transmission of Information*, M. I. T. Technology Press and John Wiley and Sons, Inc., New York (1961).

Gnedenko, B. V., *The Theory of Probability*, translated by B. D. Seckler, Chelsea Publishing Co., New York (1962).

Laning, J. H., Jr., and R. H. Battin, *Random Processes in Automatic Control*, McGraw-Hill Book Co., Inc., New York (1956).

Lee, Y. W., *Statistical Theory of Communication*, John Wiley and Sons, Inc., New York (1960).

Middleton, D., *An Introduction to Statistical Communication Theory*, McGraw-Hill Book Co., Inc., New York (1960).

Van der Ziel, A., *Noise*, Prentice-Hall, Inc., Englewood Cliffs, N. J. (1954).

Wax, N., editor, *Selected Papers on Noise and Stochastic Processes*, Dover, Publications, Inc., New York (1954).

Yaglom, A. M., *An Introduction to the Theory of Stationary Random Functions*, revised English edition, translated by R. A. Silverman, Prentice-Hall, Inc., Englewood Cliffs, N. J. (1962).

Index

Andronov, A. A., 62
Asymmetric limiter, 224
Autocorrelation function, 35

Backward equation, 64
Bateman, H., 217, 266, 285
Battin, R. H., 287
Bharucha-Reid, A. T., 64, 287

Chapman-Kolmogorov equation (*see* Smoluchowski equation)
Characteristic function, 7
  conditional, 45
Characteristic functional, 17
Coefficient of asymmetry, 10, 22
Coefficient of excess, 10, 22
Conservation of probability, equation of, 62
Correlation:
  between random variables, 10-15
  cross, 13
  double, 13
  fourth-order, 14
  multiple, 13
  triple, 13-14
Correlation coefficient, 13
  triple, 14
Correlation matrix, 40
Correlation function(s), 18
  conditional (a posteriori), 44
    of a Gaussian process, 46
  at the output of a detector, 261, 278
  of a periodic nonstationary process, 140
  of a process with stationary increments, 134
  relation to moment functions, 19
  table of, 25
  unconditional (a priori), 44
Correlation time, 22
Cramér, H., 42, 287
Cross-correlation functions, 35
Cross-spectral density, 35
Cumulants, 8
Current-voltage characteristic, 204, 243, 249, 251, 253 ff.
Cutoff frequency:
  lower, 138
  upper, 61, 136

Davenport, W. B., Jr., 193, 287
Detection of sine wave plus noise, 267-271
Detector:
  diode, 204, 243
  exponential, 253
  linear, 249
  quadratic, 251
Deutsch, R., 287
Diffusion approximation, 103
Diffusion equation (*see* Fokker-Planck equation)
Dirac delta function, 5

Edgeworth's series, 51
Envelope method, 244, 245-256
Envelope reproduction:
  coefficient, 249
  in detection, 249-254

Fano, R. M., 287
First passage, 79
First-passage time, 79
  mean, 82

problems involving, 79-83
Fluctuation equations, 84 ff.
second-order, 107-121
stochastically equivalent, 97
systems of, 100-103
Fokker-Planck equation, 62 ff.
eigenfunctions and eigenvalues of, 68
method of, in solving nonlinear problems, 263 ff.
multidimensional, 75-79
potential stationary solution of, 77
solutions of:
nonstationary, 67 ff.
stationary, 66, 109-110, 118-119, 121
two-dimensional, 107-121
Forward equation, 64
Frequency depression factor, 157

Generalized functions, 27
Gnedenko, B. V., 287
Gradstein, I. S., 72, 73, 74, 209, 231, 235, 236, 237, 241, 278, 281

Hermite polynomials, 42
generalized, 49
Hilbert transform, 183
Hille-Hardy formula, 217

Intensity coefficient(s), 22, 60, 62, 76, 86, 89

Kinetic equation (*see* Stochastic equation)
Kolmogorov equation, 64, 82
Kramers, H. A., 62
Kuznetsov, P. I., 47, 62, 149, 205

Laning, J. H., Jr., 287
Lee, Y. W., 287

Markov processes, 55-129
continuous, 62
Gaussian, 103-107
multidimensional, 75
replacement of actual processes by:
general case, 89-99, 126-129
special case, 83-89
transition from, to smoothed processes, 122-126
Method of small nonlinearity, 244, 256-262
Middleton, D., 287
Moment functions, 18
for exponential transfer characteristics, 233-242
for piecewise linear transfer characteristics, 219-232
for polynomial transfer characteristics, 209-219
relation to correlation functions, 19
Moments, 8

Narrow-band processes, 31, 36-37, 177-200
amplitude and phase of, 191
conjugate process of, 181
corresponding complex process, 180
defined by differential equations, 184-191
detection of, 245-256
Nonlinear transformations of signals and noise, 203-286
with memory (inertia), 204, 243-286
case of intermediate correlation times, 271-286
zero-memory, 203-242

Oscillatory terms, 186, 190

Poisson distribution, 144, 154
Pontryagin, L. S., 62
Potential case, 77
Probability, 4
Probability current, 62, 76
Prohability densities, 5 ff.
compatibility condition for, 16, 56
conditional, 11, 44
eigenfunction expansions of, 68-75, 110-115, 119-120
Gaussian (normal), 40
multidimensional, 10
normalization condition for, 6
at the output of a detector, 264

symmetry condition for, 16
transformation of, 6-7, 206-209
unconditional, 44
Probability functional, 17, 44
Probability integral, 42
Processes, 17 ff.
delta-correlated, 60-61, 96, 102
deterministic, 39
ergodic, 31
exponentially correlated, 24, 105
Gaussian (normal), 39-47
bivariate moments of, 212
univariate moments of, 212
Markov (*see* Markov processes)
narrow-band (*see* Narrow-band processes)
nonstationary, 131-141
periodic, 139-141
quasi-Rayleigh, 198-200
Rayleigh, 73, 192, 216-219
bivariate moments of, 218
univariate moments of, 217
slowly varying, 178 ff.
complex, 180
smoothed, 122-126
stationarily correlated, 35
stationary (*see* Stationary processes)
switched-on, 132-133
Wiener, 135, 197
with slow nonstationary changes, 131-132
with stationary increments, 133-139
without aftereffect, 55
Pulse sequences, 161-176

Quasi-moment functions, 47-53
Quasi-static approximation, 244 ff.

Random functions, 15 ff.
Random points, systems of, 143-176
correlated in pairs, 159
correlation functions of, 147
defined by first two moments, 159-161
distribution functions of, 144
conditional, 149
ergodic, 147
generating functional of, 145
nonapproaching, 160-161
Poisson, 146, 153-159
random functions constructed from, 150-153
stationary, 147
Random processes (*see* Processes)
Random spectrum, 27-31, 36
relation to spectral density, 28
Random variables, 3 ff.
completely correlated, 12
correlation between, 10-15
covariance of, 13
distribution function of, 5
mean square of, 4
mean value of, 3
standard deviation of, 9
statistically independent, 12
variance (dispersion) of, 9
conditional, 47
Rayleigh distribution, 192
Realizations, 3, 15
Relaxation time, 99
Representative point, 63
Rice's method, 233-242
Root, W. L., 193, 287
Ryshik, I. M., 72, 73, 74, 209, 231, 235, 236, 237, 241, 278, 281

Sample value, 3
Schottky's formula, 156
Seckler, B. D., 287
Semi-invariants (*see* Cumulants)
Sheppard, W. F., 42
Shot noise, 153-159
Silverman, R. A., 287
Smoluchowski equation, 57
Smoothing time, 122
Spectral densities, 21 ff.
calculation of, by use of Laplace transform, 26-27
rational, 105
table of, 25
of various pulse sequences, 161-176
Spectrum, 27
Stationary process(es), 21-37
correlation functions of, 21-22

definition of, 21
spectral densities of (*see* Spectral densities)
time averages of, 31-34
Statistical ensemble, 3
Stieltjes, T. J., 42
Stochastic equation, 59, 86, 92 ff.
Stratonovich, R. L., 47, 62, 143, 149, 205
Symmetrizing operation, 19 ff.

Tikhonov, V. I., 47, 62, 205
Transfer (dynamic) characteristics, 205 ff.
approximations to, 205
Transition probability, 56, 75, 107

Van der Ziel, A., 287
Vitt, A. A., 62

Wax, N., 287
White noise, 61
Gaussian, 61

Yaglom, A. M., 287